

**Investigating Priceless Orange Roughy (*Hoplostethus atlanticus*)
Population Dynamics using Linear Models
of Catch Per Unit Effort (CPUE)**

A thesis submitted in partial fulfilment of the requirements for the
Degree of Master of Science in Statistics

By Lisa Marie Hall
Under the supervision of
Dr. Elena Moltchanova
Dr. Alex James
and
Dr. Sharyn Goldstien

School of Mathematics and Statistics
University of Canterbury

2016

*In loving memory of Merlene,
whose support and encouragement
made a significant difference.*

SpongeBob SquarePants: Come on, Patrick. Let's do something fun.

Patrick: Would you like to run some statistics, or observe some phenomena and formulate hypotheses about said phenomena?

*- SpongeBob SquarePants
Episode: Patrick SmartPants (Smart, 2005)*

Contents

List of figures.....	6
List of tables	11
Acknowledgements.....	14
A Note for the Reader.....	15
Abstract.....	17
Acronyms	18
Introduction	19
What's the Catch?.....	19
Fisheries Science and Management in NZ	24
Fisheries Modelling Methods.....	28
Maximum Sustainable Yield (MSY)	28
Catch Per Unit Effort (CPUE)	31
Taking Stock of Orange Roughy	35
Orange Roughy Biology and Susceptibility to Overfishing.....	36
Sub-Antarctic Orange Roughy.....	38
Methodology – Fish & Statistics.....	40
Data Overview.....	40
Variables.....	40
Southern Oscillation Index (SOI)	43
Priceless Region Boundary.....	45
Defining CPUE for Priceless Orange Roughy	49
Modelling CPUE using Linear Models (LMs)	50
LM assumptions and diagnostics	51
Modelling CPUE using Temporal Variables.....	53
Segmented Linear Models and the Identification of Breakpoints	57
Model Comparisons using Akaike's Information Criterion (AIC)	59
Bootstrapping: Pulling CPUE up by its Bootstraps	61
Cross Validation: To Explain or to Predict?	63
A Note on Boxplots	65
Results of Modelling CPUE.....	66
Exploratory Data Analysis	66
Standardising CPUE using Linear Models.....	71

Standardising CPUE using Environmental and Effort Variables	77
Breakpoint Analysis.....	83
Using Bootstrapping for Pre and Post Breakpoint Comparisons	90
Model Cross Validation and the Pursuit of the Best Model for CPUE	93
Discussion.....	95
Heated Explanations of CPUE	95
And the Winner is.....	96
Biological Relevance of a Breakpoint.....	100
Practical Usage of a Breakpoint	101
Does CPUE Reflect Orange Roughy Abundance?.....	102
A Regional Weigh-in.....	104
Future Work.....	105
Incorporating Biological Parameters	105
Defining the Effort in CPUE	105
The Sub-Antarctic and Beyond.....	106
Orange roughy and oreo	106
Orange Roughy at the End of the Line	107
Bibliography	108

List of figures

Figure 1 Orange Roughy Grandpa, painted by Ellen Marcus (2011). Reproduced with permission.	16
Figure 2 A map of the New Zealand Exclusive Economic Zone (EEZ) divided into orange roughy Quota Management Areas (QMAs). NZ has one of the largest EEZ's in the world which is a contributing factor in the country becoming a world leader in fisheries management (Ministry for Primary Industries, 2014). Reproduced with permission.	23
Figure 3 An example of biomass fluctuating around the Harvest Strategy Standard (HSS) target in relation to the soft and hard limits. Targets are set with the expectation that biomass will fluctuate around them with 50% of stocks expected to be above their target and 50% below at any given time. (Ministry for Primary Industries, 2013). Reproduced with permission.	27
Figure 4 Intuitively, average catch increases as fishing effort increases. This increasing relationship has an upper bound, referred to as the Maximum Sustainable Yield (MSY). Past this point, increased fishing effort results in decreased catch and without management intervention can result in population collapse (Bonfil, 2005). Reproduced with permission.	29
Figure 5 The five stages of fishing in relation to catch and effort where MSY is the middle stage. Before MSY is reached, fishing pressure can be considered moderate or intensive, both of which support long-term sustainability. Once fishing effort results in MSY being exceeded, the population becomes overfished or severely depleted leading to a fishery which is biologically and potentially economically unsustainable (Garcia, 1996). Reproduced with permission.	30
Figure 6 The three possible relationships between CPUE and abundance are hyperstability, proportionality and hyperdepletion. Proportionality is the simplest of these, but possibly the most unrealistic in practice. The choice between hyperstability, a lagged decrease in CPUE in relation to decreases in abundance, and hyperdepletion, an accelerated decrease in CPUE as abundance decreases, is dependent on factors which include the species and fishing equipment used (Hilborn & Walters, 1992). Reproduced with permission.	32
Figure 7 An example of the possible disjoint between CPUE and abundance. The CPUE in section B-C loosely follows the abundance but the introduction of more efficient fishing gear, the discovery of a new spatial aggregation of the target species, or similar leads to a period of hyperstability in section C-D where abundance is steadily decreasing while CPUE remains steady. This disjoint is corrected in section D-E, but this alignment may not come in time to prevent the stock being overexploited (King, 2007). Reproduced with permission.	33

Figure 8 Reported worldwide orange roughy catch from 1977 to 2001. It is interesting to note the similarity between the time series of worldwide orange roughy catch (blue squares) with the MSY curve (Figure 5). The worldwide peak catch occurred in 1991, coinciding with the peak catch in Australia, but two years after the peak NZ catch in 1989. The NZ catch of orange roughy (purple cross) is greater than that of the other six countries combined (Lack, Short, & Willock, 2003). Reproduced with permission.	37
Figure 9 Total landings and TACC for ORH3B QMA from 1986 – 2013. Large landings (landed catch) occurred in the early phases of the ORH3B fishery. Both landings and the TACC decreased until a period of stability from 1995 – 2006, before another period of decrease. Catches in the Priceless region are included in the ORH3B TACC (Ministry for Primary Industries, 2014). Reproduced with permission.....	39
Figure 10 Map of the ORH3B QMA consisting of the Chatham Rise and Sub-Antarctic stocks. The Priceless, North Pukaki and Spawning Box stocks all belong to this region. The map includes bathymetric contours and the EEZ boundary (Ministry for Primary Industries, 2014) Reproduced with permission.....	47
Figure 11 Map showing the start locations of fishing events involving orange roughy and oreo catch in the NZ EEZ. Locations of catches in the Priceless, North Pukaki and Spawning Box regions are highlighted for reference. The catch locations plotted on the NZ mainland are the result of measurement errors.	48
Figure 12 Two diagnostic plots for testing the assumptions of linear models. On the left, the homoscedasticity assumption is tested by plotting the residuals against the fitted values. A good result for this plot is random scattering around the line at zero. The diagnostic plot on the right tests for normality of the residuals. The points will fall in an approximately straight line if this assumption holds.....	52
Figure 13 Simulated graphs to represent the underlying assumptions of modelling CPUE with year and month as categorical or continuous variables. Plot A demonstrates annual fluctuations in CPUE when modelled with year as a categorical variable. Plot B demonstrates monthly fluctuations in CPUE when modelled with month as a categorical variable. The fluctuations resulting from these categorical variables are unpredictable and do not follow an obvious pattern. Plot C demonstrates a decreasing linear trend in CPUE which results when year is a numeric variable. Similarly, Plot D demonstrates a monthly linear trend in CPUE.	54

Figure 14 Simulated graphs to represent the behaviour of the monthly harmonic variables and a segmented linear trend. Plot A to E demonstrate the first five monthly harmonics which include one to five peaks and troughs within each year respectively. Plot F demonstrates the behaviour of the breakpoint model. The breakpoint model occurs when there is an annual linear trend in CPUE (Figure 13C) where the slope of this trend is different either side of the breakpoint. The breakpoint model is discussed presently..... 56

Figure 15 Exploratory data analysis for catch data from the Priceless region from 1998 - 2012. Plot A is a Tukey boxplot of catch (kg) in the Priceless region plotted by year. The number of fishing events per year are plotted in a bar graph in Plot B, with the peak number of events occurring in 2004. Tonnes are used as the units of yearly catch in Plot C which is a bar graph of the total catch weight caught in the Priceless region in the most active years of the fishery. The peak catch occurred in 2002, two years prior to the peak in number of events. The components of Plots B and C are combined to create Plot D which is the annual average CPUE where CPUE is measured as kg per event. The peak annual CPUE occurred in 2002 which is the year with the peak annual catch (Plot C). 2004, the year with the peak number of fishing events (Plot B) has one of the lowest annual CPUE values of the series. There is a general decreasing trend in annual CPUE from 2005 onwards (Plot D) which is similar to the decreasing trend seen in annual catch from 2006 onwards (Plot C)..... 67

Figure 16 Exploratory data analysis for catch data from the North Pukaki region from 1989 - 2012. The time series of catch data for North Pukaki begins in 1989 with very few fishing events occurring in the exploratory phase prior to 1994. The number of fishing events has a generally increasing trend over time (Plot B) which is approximately mirrored by catch (in tonnes) (Plot C). The annual CPUE is not as promising (Plot D). The peak annual CPUE occurred in 1995 and past this point there is a generally decreasing trend in annual CPUE..... 68

Figure 17 Exploratory data analysis for catch data from the Spawning Box region in the East and South Chatham Rise from 1989 - 2014. The boxplot of catch by year demonstrates the high level of variability between individual catches (Plot A). The number of annual fishing events (Plot B) and the yearly catch (Plot C) exhibit very similar behaviour and share a peak in 2002. The average annual CPUE for the Spawning Box reaches its peak seven years later in 2009 (Plot D). There is a generally increasing trend in the annual CPUE which contrasts to the trends seen in annual CPUE in the Priceless and Pukaki regions (Figure 15D and Figure 16D)..... 69

Figure 18 Exploratory data analysis for catch data from the NZ EEZ from 1989 - 2014. The variation in catch in the NZ EEZ is greater than the variation seen in the three other regions which is expected due to the larger volume of catches included (Plot A). The peak number of fishing events occurs in 2003 (Plot B), one year after the peak in yearly catch (Plot C). The peak in annual CPUE occurred in 2014, the final year in the series (Plot D). Annual CPUE was high in the initial years and from 1994 onwards exhibits a generally increasing trend.	70
Figure 19 An exploration of sea surface temperature by location. The sea surface temperature recorded in the dataset ranges from 4.5 – 19 °C. This temperature range has been divided into three buckets, each of which span 5°C. The temperature gradient is plotted using the coordinates of the starting location of each trawl in the Priceless region. The three colours, yellow, orange and red correspond to a sea surface temperature measure in one of the three buckets, where the redder the colour, the warmer the temperature.	79
Figure 20 Time series of annual sea surface temperature measured in °C. A cyclic pattern can be seen in sea surface temperature prompting investigation into whether modelling CPUE with the SOI as an explanatory variable could lead to further insights and greater accuracy.....	80
Figure 21 A comparison of CPUE for El Niño, La Niña and neutral conditions. The El Niño and neutral conditions have a similar effect on CPUE, whereas the La Niña conditions result in a lower CPUE overall.....	82
Figure 22 CPUE plotted by year with an apparent breakpoint around 2004. This suggests that an LM that includes an annual linear trend that incorporates a breakpoint could lead to an improved explanation of CPUE.....	83
Figure 23 A plot of AIC values for candidate breakpoint years for modelling CPUE with a segmented linear trend. The breakpoint models are LMs fitted with an annual linear trend in CPUE where the slope of the trend is allowed to differ either side of the breakpoint. Fifteen separate models are fitted with the candidate breakpoint ranging from 1998 to 2012. The minimum AIC leads to identification of the year 2004 as the optimal breakpoint.	85
Figure 24 A plot of trawl locations either side of the 2004 CPUE breakpoint. Start and end locations of fishing events have been plotted and joined with directed line segments. The clustering of trawl locations enables identification of the spatial features of the region such as valleys and seamounts. These features correspond to spatial distributions of orange roughy aggregations. The trawl locations have a similar spatial distribution either side of the breakpoint which rules out location as a possible explanation for the breakpoint.....	92

Figure 25 A plot of the number of events in the Priceless region by month over a period of 180 months from January 1998 – December 2012. October is the most popular month for fishing events due to the fact that the fishing year starts on the 1st of October and this corresponds to the start date for the ACEs. 98

Figure 26 A plot of the fitted vs observed annual CPUE. The fitted model for the best LM of CPUE is plotted (solid red line) with 95% confidence intervals either side (dashed red lines). The fitted model is overlaid on the observed annual CPUE data (empty grey dots) and a contour plot of the number of events per year sits on the x axis to give an indication of relative sample sizes (solid grey contour). The fitted model clearly demonstrates the location of the breakpoint in 2004 and the differing slope either side of this year. 99

List of tables

Table 1 Harvest Strategy Standard measures and definitions (Ministry for Primary Industries, 2013)	26
Table 2 Timeline of development of the orange roughy fishery in the Sub-Antarctic from 1990 to 2010 (Ministry for Primary Industries, 2010)	38
Table 3 Summary of useful commercial catch variables for the Priceless region used for exploratory analysis and modelling. These variables are from the fishing event and estimated catch datasets...	41
Table 4 Definition of the Priceless region with coordinates in degrees, minutes and seconds (DDD, MM, SS) and Decimal Degrees (DDD.DDDD) (Anderson & Dunn, 2012) (CSGNetwork, 2016)	45
Table 5 Definition of the North Pukaki region, located in the Sub - Antarctic, with coordinates in degrees, minutes and seconds (DDD, MM, SS) and Decimal Degrees (DDD.DDDD). The North Pukaki region is a rectangle defined by the coordinates in the table excluding the area defined as the Priceless region (Anderson & Dunn, 2012) (CSGNetwork, 2016)	46
Table 6 Definition of the Spawning Box region, located within the East and South Chatham Rise, with coordinates in degrees, minutes and seconds (DDD, MM, SS) and Decimal Degrees (DDD.DDDD). The Spawning Box is one of the key orange roughy stocks located in the East and South Chatham Rise (Ministry of Fisheries, 2008) (CSGNetwork, 2016)	46
Table 7 R syntax for including explanatory variables in a linear model.	51
Table 8 Formulas for the monthly harmonic variables	55
Table 9 AIC values and model interpretations for CPUE models with variants on year and month as explanatory variables. The Δ AIC and AIC weights are included in brackets to the right of the model AIC. The model with the lowest AIC is highlighted in bold and implies abrupt changes in monthly and annual CPUE.	72
Table 10 AIC values for CPUE models with variants on year and month and their interactions as explanatory variables. The Δ AIC and AIC weights are included in brackets to the right of the model AIC. The model with the lowest AIC is in bold and models CPUE using month with three seasonal harmonics which vary annually and year with abrupt changes.	73
Table 11 AIC values for CPUE models with species and variants on year and month as explanatory variables. The Δ AIC and AIC weights are included in brackets to the right of the model AIC. The model with the lowest AIC is in bold and includes species and year and month as categorical variables. This model implies that a change in species, year or month corresponds to an abrupt change in CPUE.	74

Table 12 AIC values for CPUE models with species and variants on year and month as explanatory variables including interactions. The Δ AIC and AIC weights are included in brackets to the right of the model AIC. The model with the lowest AIC is in bold and includes species, month with three harmonics, categorical year and the full set of interactions.	75
Table 13 AIC values for CPUE models with a selection of effort and environmental variables. The adjacent columns include the Δ AIC and AIC weights to facilitate comparison of model performance. The best model is in bold and includes species, month with three harmonics, year as a categorical variable, surface temperature and the corresponding interactions between these variables.	77
Table 14 AIC values for models of CPUE using the SOI and a categorical variable corresponding to El Niño conditions as explanatory variables. These two variables have been added to the best temporal model consisting of species, month with three harmonics and abrupt changes in in yearly CPUE (Table 12). This model and the model including surface temperature have been included in the table for comparison. The Δ AIC and AIC weights indicate that surface temperature cannot be adequately replaced by either the SOI or El Niño variables.	81
Table 15 AIC values for models of CPUE which incorporate a linear trend with a breakpoint in 2004 and variants of month as explanatory variables. Variable interactions and species are not included in these models. The Δ AIC and AIC weights are included in brackets to the right of the model AIC. The best of these breakpoint models is in bold and includes abrupt monthly fluctuations in CPUE...	86
Table 16 AIC values for models of CPUE which incorporate a linear trend with a breakpoint in 2004 and variants of month as explanatory variables including interactions. The Δ AIC and AIC weights are included in brackets to the right of the model AIC. The model with the lowest AIC is in bold and includes abrupt monthly changes in CPUE. The model including month with four harmonics has the second lowest AIC with a weight of 0.25 making it a candidate for the best model in this table.	87
Table 17 AIC values for models of CPUE which incorporate a linear trend with a breakpoint in 2004 and variants of month and species as explanatory variables. The Δ AIC and AIC weights are included in brackets to the right of the model AIC. The model with the lowest AIC is in bold and includes abrupt monthly changes in CPUE. The models including month with three, four or five harmonics also have non-zero weights.	88
Table 18 AIC values for models of CPUE which incorporate a linear trend with a breakpoint in 2004 and variants of month and species as explanatory variables including interactions. The Δ AIC and AIC weights are included in brackets to the right of the model AIC. The model with the lowest AIC is in bold and includes abrupt monthly changes in CPUE. The models including month with four or five harmonics also have non-zero AIC weights.	89

Table 19 A comparison of the means of a selection of relevant explanatory variables pre and post the 2004 breakpoint including 95% confidence intervals and the corresponding p-values from the independent two sample t-tests. This comparison was performed using bootstrapping with a bootstrap sample of 1000. The variable units and degree of rounding for each of the means and confidence intervals are given alongside the variable name in the first column.	90
Table 20 Summary of results from 10-fold cross validation of the key LMs for CPUE. The errors calculated as part of the cross validation are the MAE and the RMSE. The table has been arranged with the MAE in descending order. Two of the RMSEs are larger than those below them in the table which indicates increased variance of the predictions. These two errors are underlined. The inclusion of model AIC values allows for an overall comparison of model performance. Results for models including month as a categorical variable (month_F) should be treated with caution due to issues with rank deficiency. The overall best model for CPUE can be found in bold and includes species, month with three harmonics and an annual linear trend with a breakpoint in 2004.	94

Acknowledgements

Over the past year I have had a great team of mentors to help bring this thesis together.

Elena, Alex and Sharyn – thank you for all the time you have invested in this project. Thank you for sharing your advice and experiences on how to do scientific research well. My understanding of statistics, coding, scientific process and scientific communication has undergone vast improvement as a result of your input.

Thank you for your patience in explaining the same concepts to me week after week until I finally clicked, for encouraging me to engage in seminars and current literature, for constantly challenging my understanding, for your proof-reading and advice on how to make this a better a thesis, and ultimately, for guiding me through months of feeling out of my depth, to a thesis I am proud of. I am grateful for your support and intellectual input.

To Alistair Dunn from NIWA – thank you for helping me get started with this project and for your advice on getting to grips with fisheries data.

To Rich Ford and Pamela Mace from MPI – thank you for the opportunity to do this project and for your patience and understanding when I needed to take some time out.

To Matt Dunn – thank you for your assistance in writing my thesis proposal and giving me ideas on where this project could take me.

And finally, a big thank you to MPI and NIWA for funding this research and providing the data.

A Note for the Reader

As I have been working on this thesis, one of the common questions I have received from friends and colleagues is whether or not I think orange roughy are a sustainable catch. I will freely admit that I have eaten orange roughy – I bought a box of frozen fillets a month before starting this project. I'm still undecided as to whether I would eat orange roughy again. I thought that by now I might have an answer, but the deeper I've gone with my research, the murkier the waters have become.

Two thirds of the way into this thesis, I came upon the realisation that I had no idea what sustainability is. Given the topic of my thesis and the fact that fisheries management is conducted with sustainability in mind, it was imperative that I rectified the situation.

What on earth is sustainability? Sustainability has become a buzzword, not only in fisheries, but in any field where the use of the word has the potential to improve public perception. Its use is as ubiquitous and inconsequential as mention of the Kardashians. Stereotypical images of the world in the palm of a hand are emotive, but their interpretation is vague. As a result, the word sustainability requires quantification in a given context in order to determine its precise meaning. A clear definition is especially important when evaluating the success or failure of a sustainability initiative.

The work of Daniel Pauly and Ray Hilborn, two of the most well-known and controversial fisheries scientists in the world, has played a big part in shaping my views of the oceans. Ray Hilborn is well-known for fierce fights and critiques of Daniel Pauly (Cressey, 2015). The first fisheries textbook I picked up was one co-written by Hilborn (Hilborn & Walters, 1992) so I started out on Team Hilborn. Throughout the last 12 months, I have swung like a chaotic pendulum between the two opposing views towards fisheries science. The debate published in *Nature* between Pauly, Hilborn and Branch on using catch data to reflect abundance swayed me to Team Pauly for a significant chunk of the year (Pauly, Hilborn, & Branch, 2013). Now the pendulum has reached equilibrium and I have adopted ideas from both sides of the debate.

In weeks filled with optimistic readings of fisheries management I didn't think twice about eating fish, but some weeks I felt obliged to interchange my fish with a hotdog. Should you continue to eat fish? That's not for me to decide. However, I have come to realise that the choice to eat fish is as much a moral dilemma as an environmental one. I believe that everyone should know where their food has come from and its potential environmental impact from the perspective of both the managers and conservationists. If that information isn't enough to put you off your dinner, then I wish you *bon appétit*.

Should you continue to eat orange roughy? To answer this question for yourself, refer to Figure 1, the Orange Roughy Grandpa. This picture provides a more poignant answer to that question than a thousand of my words ever could.

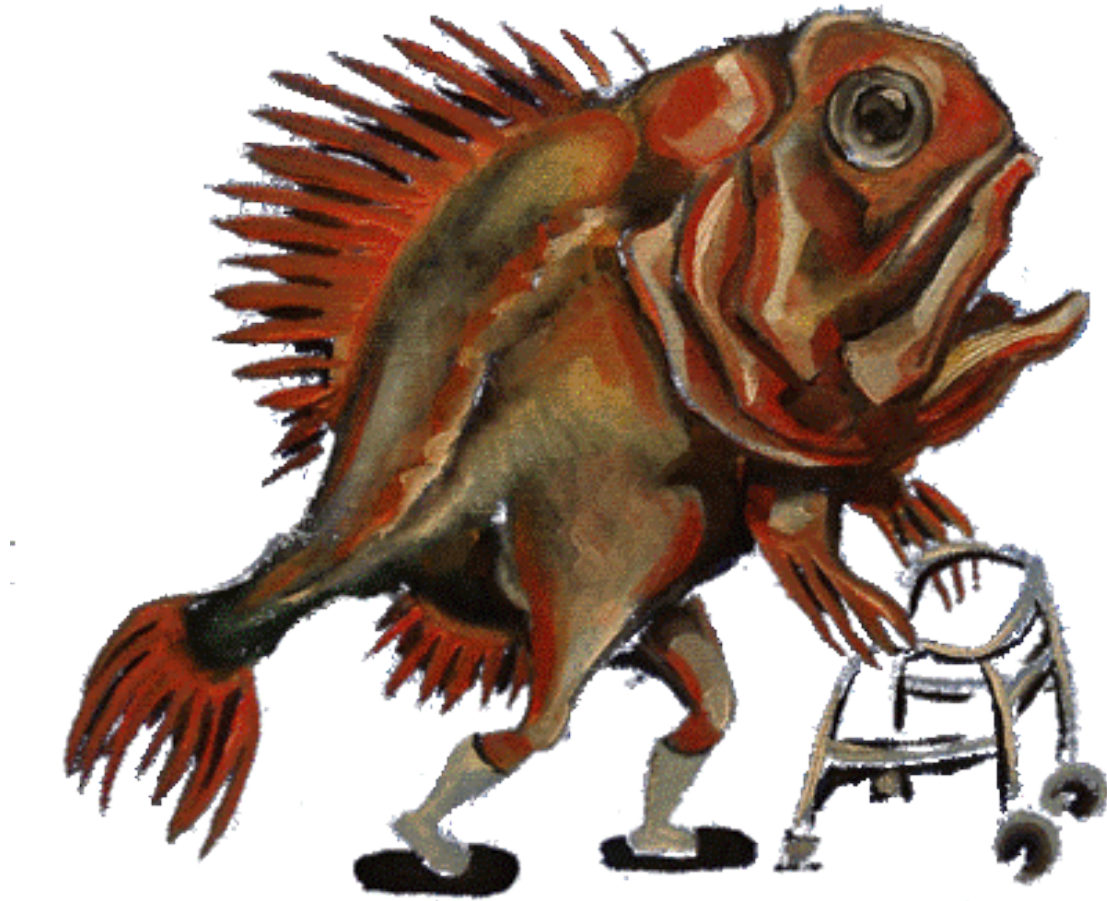


Figure 1 Orange Roughy Grandpa, painted by Ellen Marcus (2011). Reproduced with permission.

Abstract

Who do you trust when trying to make an informed and environmentally friendly choice on fish fillets at the supermarket? The Marine Stewardship Council and Forest and Bird have contradictory opinions when it comes to orange roughy (*Hoplostethus atlanticus*) sustainability. The controversy ultimately comes down to the fact that orange roughy is a deep sea species and while management may have been successful for inshore species, the vast number of unknowns relating to deep sea stocks and the ecosystem effects of deep sea fishing mean that what initially seemed like sustainable management repeatedly resulted in disastrous consequences. The collapse of some of the early orange roughy stocks is proof of this. The question is, have we learnt enough from those early failures to avoid repeating those mistakes, or is it only a matter of time before the next crash of one of our deep sea fisheries?

Fitting a selection of linear models to catch per unit effort (CPUE) data, corresponding to orange roughy catch from the Priceless region in the Sub-Antarctic, led to the discovery of significant seasonal patterns and an annual linear trend in CPUE with a breakpoint in 2004.

The assumed relationship between CPUE and abundance is difficult to quantify but the cessation of fishing in the region leads to the conclusion that decreases in CPUE correspond to decreases in orange roughy abundance.

Acronyms

Included below is a list of the acronyms which have been used throughout this report.

ACE: Annual Catch Entitlement

CPUE: Catch per Unit Effort

EEZ: Exclusive Economic Zone

GLM: Generalised linear model

HSS: Harvest Strategy Standard

ITQ: Individual Transferable Quota

LM: linear model

MAE: Mean absolute error

MPA: Marine Protected Area

MPI: Ministry for Primary Industries

MSC: Marine Stewardship Council

MSE: Mean squared error

NIWA: National Institute of Water and Air

NZ: New Zealand

ORH: Orange roughy (*Hoplostethus atlanticus*)

OEO: Oreo (*Allocyttus niger*, *Allocyttus verucosus*, *Neocyttus rhomboidalis* and *Pseudocyttus maculatus*)

QMA: Quota Management Area

QMS: Quota Management System

RMSE: Root mean squared error

SLP: Sea Level Pressure

SOI: Southern Oscillation Index

TACC: Total Allowable Commercial Catch

Introduction

What's the Catch?

'We account the whale immortal in his species, however perishable in his individuality' (*Melville, 1851*).

This quote from the famous novel *Moby-Dick* epitomises the attitude towards much of the life in the oceans. Time and time again the perishability of a species has not been considered until it can no longer be ignored.

Commercial fishing has existed for thousands of years (Love, 2010). Today the fishing industry plays a significant role in global trade and nutrition with global exports of fish totalling US\$129.2 billion and annual world per capita consumption of fish averaging 19.2kg in 2012 (FAO, 2014). Centuries of increased fishing pressure have resulted in widespread losses of biodiversity and species abundance. The catch of this increased pressure is captured in the prediction that global collapse of all currently exploited taxa will occur in the year 2048 in the absence of an overhaul of current fisheries management (Worm, et al., 2006). This prediction made world headlines and stirred up considerable debate amongst the fisheries science community.

Some fisheries scientists, including Daniel Pauly, began using the term *aquacalypse* to describe the state of the world's fisheries. In his article entitled *Aquacalypse Now*, Pauly references the 2048 prediction by expressing uncertainty about the accuracy of the year or the decade of fisheries collapse but concluding that irrespective of the prediction accuracy, fish are in dire peril (Pauly, 2009). Pauly's own aquacalyptic predictions deal with the fishing down of the ocean food web with the potential outcome of an ocean overrun by inedible jellyfish (Pauly, 2010).

Ray Hilborn has spent much of his career advising governments on fisheries management strategies. His portfolio includes many fisheries management successes so it came as no surprise that he was ready to rebut the aquacalyptic claims of the early 2000s. Fisheries scientists often have inconsistent views when it comes to defining management outcomes as a success or failure. This conflict is often a result of focus on abundance rather than sustainable yield. A fish stock at relatively low population levels in comparison to a historic baseline while still producing at its maximum sustainable yield may indicate management failure to those with a focus on abundance, while simultaneously indicating management success to those with a focus on sustainable yield. An example of one such management success is the reduction in the proportion of overfished stocks from 33% to 26% between 2001 and 2005 in the U.S.A. (Hilborn, 2007).

While Worm et al. (2006) discussed the limitations of their prediction, Hilborn criticised it for being an exaggerated warning and affirmed that fish do not need to be taken off the menu (Hilborn, 2011). This healthy scientific debate led to a collaboration between Worm, Hilborn and a selection of other fisheries experts from around the world. This collaboration used prediction methods that both parties agreed upon and led to a more optimistic view than Worm et al.'s 2006 paper without denying the large scale issue of overexploitation (Worm, et al., 2009).

The first and second Great Fishing Experiments refer to periods during the first and second World Wars. During these periods the level of fishing around Europe was significantly reduced. In Britain the effort from their trawl fleet was reduced by 97%, resulting in a temporary Marine Protected Area (MPA) (Malakoff, 2010). A group of scientists at the Lowestoft Fisheries Laboratory analysed the effects of both fishing experiments and found that the density of many species increased three fold due to increased survival to older ages. Many were finding that catch rates had increased post-war and it soon became apparent that something needed to be done to address the imminent threat of overfishing. This led to the International Overfishing Conference in 1946 (Smith, 1988). After this conference, fishing regulations and the use of statistical and mathematical models became more and more common within the industry. The theory behind temporary fishery closures and MPAs was given an unexpected boost and the lessons learnt from post-war fisheries continue to inform current research.

Discussion around the management of fisheries resources is not a new phenomenon. Sumerian and Babylonian writings dating from 5000 BC include reference to fishing rights (Love, 2010). Early in the 17th century the legal concept of freedom of the seas was developed. The basis of this concept was that the oceans were an infinite resource with unrestricted access for transport or exploitation of resources. Water within three nautical miles of land was considered territorial sea and under the jurisdiction of the adjacent country (Mansfield, 2012). Waters outside of the territorial sea were subject to unregulated fishing pressure. At this stage there were plenty of fish to go around so ensuring long-term sustainability was not on the agenda.

In 1883 Thomas Huxley famously reasoned that sea fisheries were probably inexhaustible. He felt that nothing the fishermen were doing seriously affected the number of fish and that attempts to regulate fisheries would prove useless. Huxley's assumptions were based on fishing levels and technology of the time so are not quite as ridiculous as they appear in retrospect. However, his observation that the multitude of cod, herring and mackerel was inconceivably great, highlights the detrimental impact of overestimating the population of fish stocks. Similar population overestimates have been repeated in fisheries around the world, including New Zealand's orange roughy stocks.

Huxley was surprised to discover that most fishermen knew nothing about fish, except how to catch them. While improvements in fisheries science now mean that much more is known about the stocks being exploited, many of the mistakes from the past continue to be repeated as the biological, economic and social components of fisheries management compete for precedence. Upon the discovery of a new stock there is often insufficient time and money to develop an understanding of the biology or reproductive history of the species which has led to overexploitation in many fish stocks. Species which are long-lived and have low fecundity have a higher risk of overexploitation.

Not everyone shared Huxley's optimistic view of fisheries. Lankester was one of those present at the Fisheries Exhibition in 1883 and challenged Huxley's view of inexhaustible fisheries and surplus fish. He argued that the removal of any fish had an impact, either directly or indirectly by the removal of potential parents (Smith, 1988). These different views started an argument within the scientific community that lasted for decades.

While Huxley's comments continue to spark criticism, his view of the inexhaustibility of the sea has been shared by countless others, many of whom have expressed this view in more recent times. Despite evidence of overfishing and the lessons learnt from the Great Fisheries Experiments, the book *The Inexhaustible Sea* was published in 1955 by two academics in the field in which they claimed the following:

We are already beginning to understand that what [the ocean] has to offer extends beyond the limits of our imagination – that someday men will learn that in its bounty the sea is inexhaustible (Daniel & Minot, 1955).

Are fisheries exhaustible, and if so, can anything be done to prevent their exhaustion? The collapse of many of the world's important fisheries including the Atlantic cod, North Sea herring and Peruvian anchoveta fisheries, amongst others shows that fisheries are indeed exhaustible (Hilborn & Walters, 1992). There are many factors affecting fisheries which are outside our control, but there are steps which can be taken to significantly reduce the probability of exhaustion.

Given drastic reductions in cod abundance, Cape Cod in Massachusetts, U.S.A is a modern day misnomer. The area was named by Bartholomew Gosnold in the early 17th century after his amazement at the sheer abundance of cod in the area. A century later, the lack of cod abundance was the cause of amazement and resulted in fishing moving further offshore to maintain their catches (Schrope, 2010).

Fisheries are subject to what Arnason (2009) refers to as the fisheries problem. The fisheries problem is a deep-rooted problem of economic inefficiency resulting from inadequate control over fisheries resources in the form of the common property arrangement. This arrangement inevitably results in a free-for-all where fishermen compete amongst themselves for as large a share of the catch as possible. This is referred to as an Olympic-style fishery or the race for fish. This Olympic mentality removes incentives to avoid overexploitation and commonly leads to depleted stocks and higher expenses as effort increases and catch decreases.

One of the most catastrophic fisheries collapses in recent times has been linked to the race for fish. After decades of sustainability concerns, the Peruvian anchoveta fishery collapsed in 1972. The cause of collapse was attributed to overfishing coupled with very strong El Niño conditions (Aranda, 2009). A decade later, the strong El Niño conditions in 1982-83 led to a more extreme population collapse, leaving the fishery economically unsustainable. Peruvian anchoveta is the most heavily exploited fish in world history (Iwamoto, Eschmeyer, & Alvardo, 2010). Despite the biological and economic collapse of the fishery, proposed changes to management of the fishery encountered resistance until June 2008 when the government introduced rights based management to the fishery.

The introduction of rights based management, such as an Individual Transferable Quota (ITQ) system, is usually faced with strong opposition from stakeholders and politicians (Aranda, 2009). As such, many fisheries have made the shift towards rights based management out of necessity. Successful implementation requires a shift in focus from short-term gains to long-term sustainability. The advantage of a rights based management system such as ITQ, is that fisherman are given an economic incentive to look after fish stocks. The value of quotas increases if the fishery is in good condition which leads to improvements in industry performance and sustainability through changes in incentives and ultimately, the behaviour of fishermen (Stewart & Callagher, 2011). The ability of ITQs to change the mindset of fishermen is a key component of its success.

Rights based management systems have been successfully implemented in many areas around the world. The most notable countries using rights based management are New Zealand, Australia, Iceland, the U.S.A. and Canada. Between these five countries, approximately 180 species are managed under a rights based system (Sanchirico & Wilen, 2007).

One of the main criticisms of ITQs is their bias towards economic sustainability (Pauly, 2009). Research has shown that ITQs have the potential to not only slow the trend towards global fisheries collapse, but to reverse it (Costello, Gaines, & Lynham, 2008). Many of the ITQ success stories point not only to economic sustainability as suggested by Pauly, but also towards biological sustainability.

Sanchirico & Wilen (2007) sum up the situation with the following quote:

It is widely believed and supported by anecdotal evidence that once fishers have a financial stake in the returns from sensible investment in sustainable practices, they are more easily convinced to make sacrifices required to rebuild and sustain fisheries at high levels of economic and biological productivity.

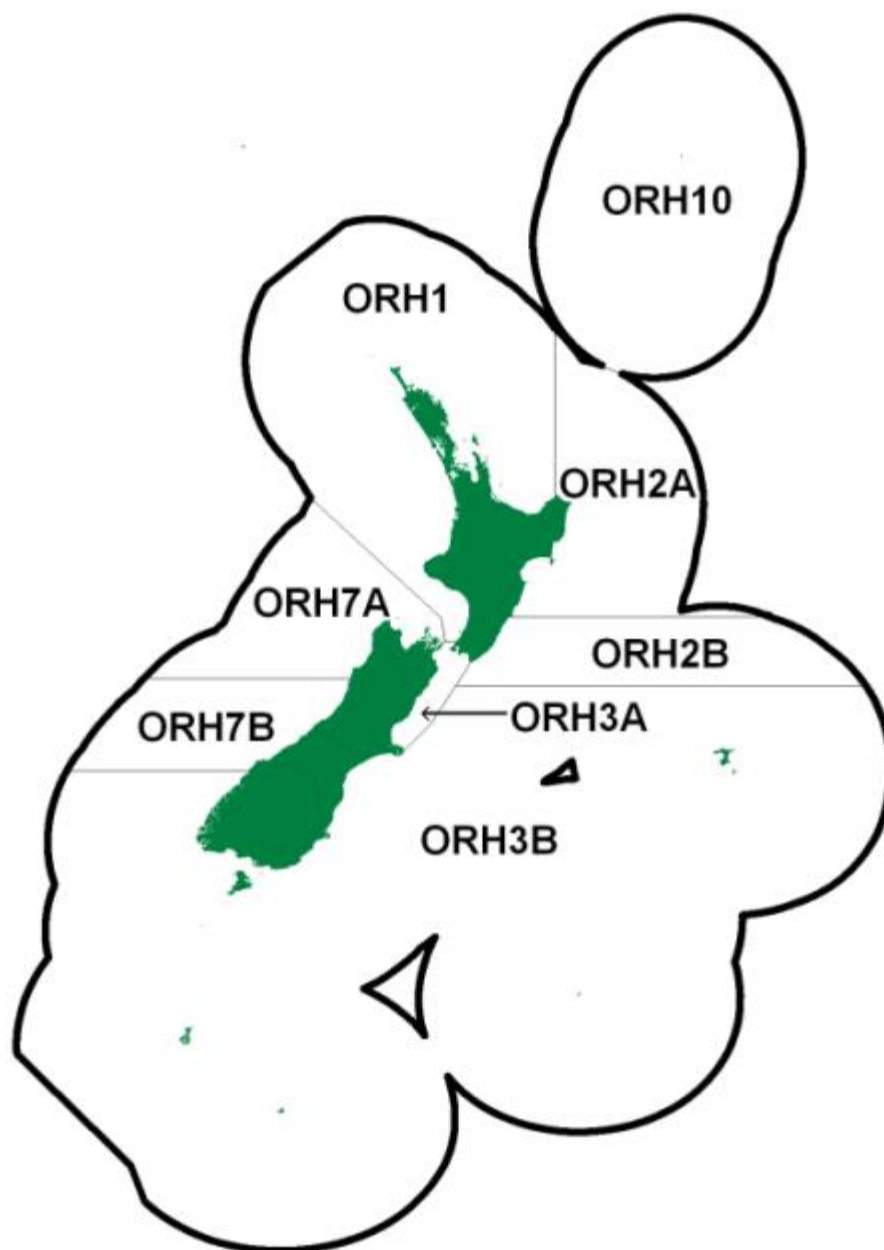


Figure 2 A map of the New Zealand Exclusive Economic Zone (EEZ) divided into orange roughy Quota Management Areas (QMAs). NZ has one of the largest EEZ's in the world which is a contributing factor in the country becoming a world leader in fisheries management (Ministry for Primary Industries, 2014). Reproduced with permission.

Fisheries Science and Management in NZ

The use of NZ's fisheries resources has wide ranging effects in the environmental, economic and social sectors. The key objective of fisheries science in NZ is thus to “maximise the value New Zealanders obtain through the sustainable use of fisheries resources and the protection of the aquatic environment” (OECD, 2015). There are three components which accompany this objective:

1. Protecting the health of the aquatic environment
2. Enabling people to get the best value from the sustainable and efficient use of fisheries
3. Ensuring obligations to Māori are met

Ensuring sustainability of fisheries resources refers to the maintenance of the potential of fisheries resources to meet the foreseeable needs of future generations and to avoid, remedy or mitigate any adverse effects of fishing on the aquatic environment (New Zealand Government, 2016).

Until 1970, New Zealand's territorial waters extended only three nautical miles offshore in accordance with the international concept of freedom of the seas. In 1970 the territorial waters were extended to 12 nautical miles until 1978 when the 200 nautical mile Exclusive Economic Zone (EEZ) was established (Walrond, 2012). The international laws behind the implementation of EEZs were implemented in 1982 at the third United Nations Conference on the Law of the Sea after a ten year negotiation period. The EEZ gives the coastal state sovereign rights to the waters within this boundary. These rights are for the purpose of exploring, exploiting, conserving and managing both living and non-living resources of the waters and the seabed, including subsoil (United Nations, 2013).

As an island nation with a territory that includes many outlying islands, including the Chatham Islands, New Zealand has one of the largest EEZs in the world with a size of over 4 million square kilometres and an average depth of 2.4 kilometres. This corresponds to around one million Olympic swimming pools worth of EEZ water per person living in New Zealand (Stevens & O'Callaghan, 2015). Depending on the EEZ boundary definition used, the ranking for the size of the NZ EEZ ranges from 4th to 9th largest (Mansfield, 2012).

Until the introduction of the EEZ legislation it was impossible for the New Zealand government to have any control over the sustainable harvesting of fisheries in our surrounding waters. The pre-EEZ period involved an Olympic race between international and local fishing companies who were competing for the best and biggest catches. Foreign companies began exploiting fisheries resources around New Zealand in the 1960s. Post-EEZ implementation, it became clear that management schemes would be necessary to protect our marine environment and ensure sustainability of the fishing industry.

New Zealand began to experience the collapse of marine resources after the stripping of oyster beds for export in the 1880s led to a collapse in the oyster fishery in the 1890s (Walrond, 2012). The fishing pressure on coastal resources in the 1980s led to them becoming fully exploited. Estimates made in 1983 suggested that the number of inshore fishing vessels needed to be reduced by around 300 to balance the available catch with the fishing effort expended. After decades of plentiful fish, the view of finfish stocks echoed that of Huxley's view of the inexhaustibility of the sea until large decreases in catch disproved this. The disparity between catch and vessel numbers led to two significant changes to the industry. The first was a move from inshore to deep-sea fisheries which included hoki and orange roughy. The second was the introduction of the Quota Management System (QMS).

The New Zealand government introduced the QMS in October 1986. Prior to the QMS, NZ fisheries were managed by controls on vessel numbers. This strategy had no way of controlling the numbers of fish caught. The QMS introduced restrictions on catch by allocating quotas to permit holders. The allocation of quotas was based on catch history. This resulted in increased fishing in the pre-implementation period as fishermen attempted to increase their quota entitlements.

One of the main building blocks of NZ's QMS is the Individual Transferable Quota (ITQ). NZ is considered to be a world leader in the use of ITQs for fisheries management (Lock & Leslie, 2007) and became the first country to apply them on a national scale (Sissenwine & Mace, 1992). Initially awarded to fishermen based on historic catch data, they are treated as a property right and have many associated benefits which include exclusivity, transferability, divisibility, flexibility and security.

There are 628 stocks in the NZ QMS, 350 of which are considered to be commercially valuable (Ministry for Primary Industries, 2013). These stocks are made up of 95 species or species groups. MPI is responsible for annually setting quotas for the commercially important species based on stock assessments which are carried out by NIWA. Each year MPI announces the Total Allowable Commercial Catch (TACC) for each stock in the QMS. This TACC is then divided amongst fishermen who hold an ITQ for that stock in the form of an Annual Catch Entitlement (ACE). The ACE is a catch right created via quota shares and can be owned in quantities of one kilogram and above. The ACE was introduced to the QMS in 2001 (Stewart & Callagher, 2011). Catch over and above the ACE requires payment of a deemed value or purchase of additional ACE for the season to cover the surplus catch. A deemed value is a fine set to discourage catches exceeding the ACE (OECD, 2015). It is set at a level to minimise discards at sea while also discouraging fishers from targeting fish they don't hold quota for.

Initial implementation of the ITQ system in NZ involved fixed ITQs where the government would buy or sell quota to fishers to adjust the TACC to a suitable level. This system was changed to variable ITQs, which entitled quota holders to a percentage of the TACC rather than a fixed quota, when it was realised that orange roughy abundance had been grossly overestimated and consequently, large reductions in the Chatham Rise orange roughy TACC were required. Under the fixed ITQ system, this reduction had the potential to cost the government over \$100 million to buy back enough quota from fishers to reduce the TACC to a level which supported long-term sustainability of the stock (Sissenwine & Mace, 1992). After changing from fixed to variable ITQs, the Chatham Rise orange roughy quota was reduced to 5000 tonnes per year, with the TACC being recalculated periodically based on new data.

HSS measure	Definition
Soft limit	A biomass level below which the stock is considered overfished and requires rebuilding
Hard limit	A biomass level below which the stock is considered collapsed and may require closure of the fishery to rebuild the stock
Overfishing threshold	A rate of fishing which leads to stock declining below management targets when exceeded
Management target	A biomass level that stocks are expected to fluctuate around with at least 50% chance of achieving the target

Table 1 Harvest Strategy Standard measures and definitions (Ministry for Primary Industries, 2013)

The Harvest Strategy Standard (HSS) is a guide for stock assessments. It is based on four measures: soft limits, hard limits, overfishing thresholds and management targets.

Given that the management target is set so that the probability of achieving the target is at least 50%, each year it is expected that around 50% of stocks will be above their management target and 50% below. Having approximately 50% of stocks below their management target at any one time has often caused alarm amongst the media and conservation groups. However, the management targets are designed to take these fluctuations into account. There is only cause for concern when, as well as being below the management target, the biomass is reduced to the soft limit. Biomass estimates are expected to fluctuate around the management target in relation to the soft and hard limits (Figure 3).

When evaluating the management strategy, there are three cases which should be considered:

1. The probability of being above the target biomass should be at least 50%.
2. The probability of the biomass being below the soft limit should be less than 10%.
3. The probability of the biomass being below the hard limit should be less than 2%.

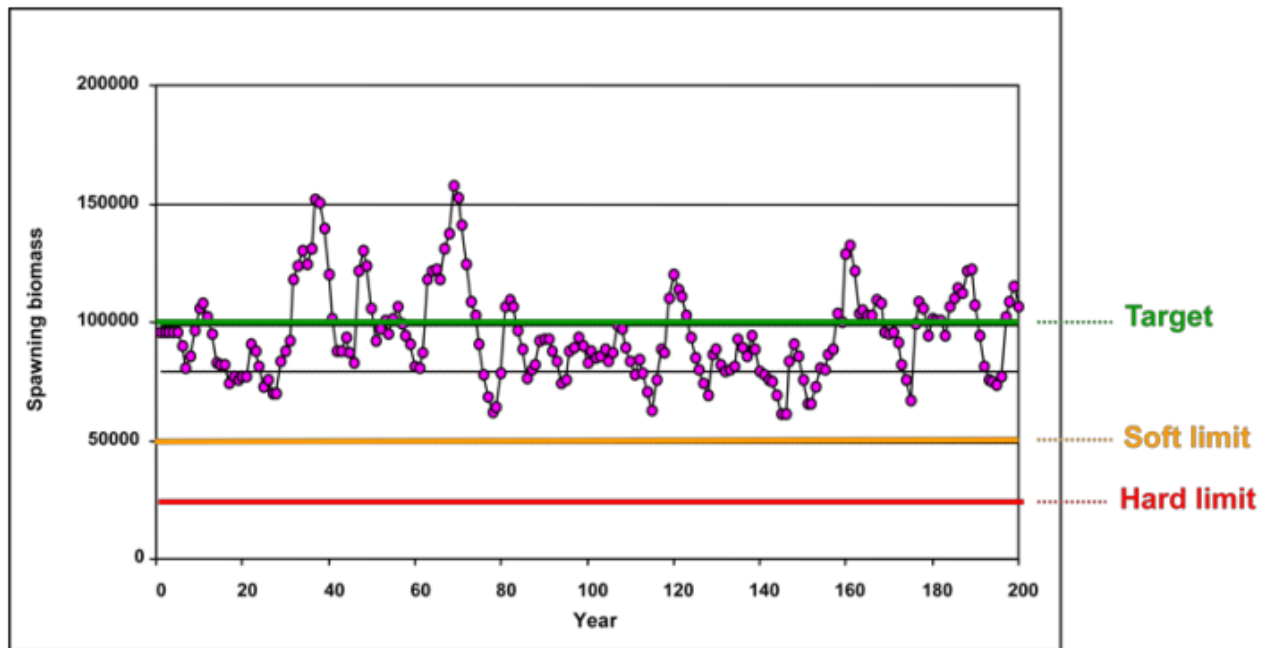


Figure 3 An example of biomass fluctuating around the Harvest Strategy Standard (HSS) target in relation to the soft and hard limits. Targets are set with the expectation that biomass will fluctuate around them with 50% of stocks expected to be above their target and 50% below at any given time. (Ministry for Primary Industries, 2013). Reproduced with permission.

The soft and hard limit cases can also be combined with the aim that the probability of the biomass being below the soft limit is less than 5% (Ministry of Fisheries, 2008). The HSS is designed to facilitate management actions which promote achieving targets and avoiding limits.

Fisheries Modelling Methods

Maximum Sustainable Yield (MSY)

There are three ways to get more from the sea when the rate of fishing exceeds the rate of production of a species – catch something else, fish somewhere else, or fish less – Callum Roberts (2007).

Following his quote, Robert recognises that getting more fish by fishing less seems paradoxical. The biological consequence of fishing less is that more fish have the potential to grow to a larger size and the observation of this phenomenon has provided a basis for modern fisheries science

One of the most common concepts in fisheries science is that of maximum sustainable yield (MSY). The basic concept is that as fishing effort increases, the catch also increases. However, this increase in catch only occurs up to a point and this point is known as the MSY. Increased effort once this point has been reached results in decreased catch as the rate of reproduction fails to compete with fishing mortality (Figure 4). Hilborn and Walters (1992) refer to the MSY curve as being perhaps the most commonly printed illustration in fisheries textbooks, and the most dangerous.

The concept of MSY has three associated assumptions:

1. There must be fishing effort in order for there to be catch.
2. When fishing effort is very high, the stock will be fished to low numbers resulting in reduced reproductive capacity and limited surplus population.
3. The maximum average yield thus occurs somewhere between no effort and very high effort.

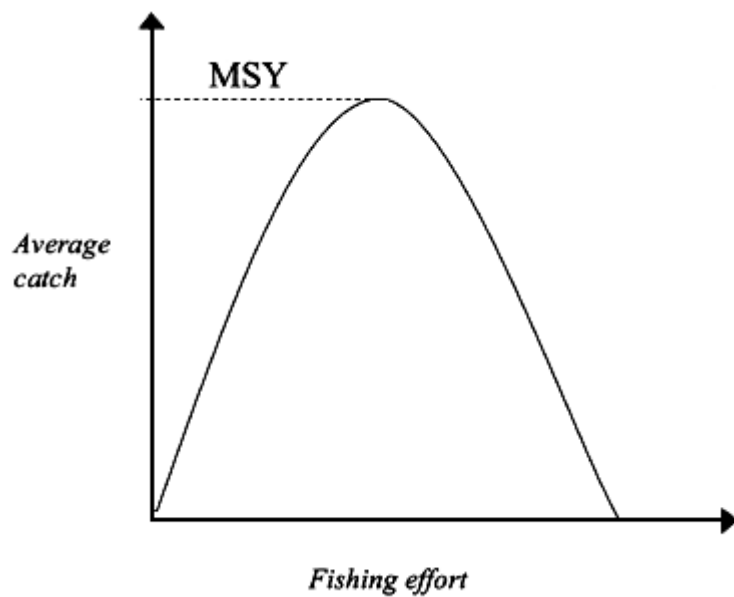


Figure 4 Intuitively, average catch increases as fishing effort increases. This increasing relationship has an upper bound, referred to as the Maximum Sustainable Yield (MSY). Past this point, increased fishing effort results in decreased catch and without management intervention can result in population collapse (Bonfil, 2005). Reproduced with permission.

The overall concept of MSY and the corresponding curve (Figure 4) seem both intuitive and simple, which begs the question - what makes MSY so dangerous? Hilborn & Walters (1992) point out that MSY often leads to asking the wrong questions. Many fisheries stock assessments choose to focus on determining the optimum effort and the MSY. Despite the simplicity of the concept, it can be exceedingly difficult to identify these quantities from data, especially when there are high levels of noise present. Theoretically, once the average catch starts to drop, identification of the MSY and associated effort should be simple. Commercial catch data is inherently noisy and can experience fluctuations for many reasons which include seasonality, and changes in technology or fishing expertise. Thus, in order to have any level of certainty around an estimate of MSY, the MSY must be exceeded, sometimes to the point of overfishing (Figure 5). Going too far beyond the MSY can lead to a fishery which is depleted, leading to concerns over biological and economic sustainability.

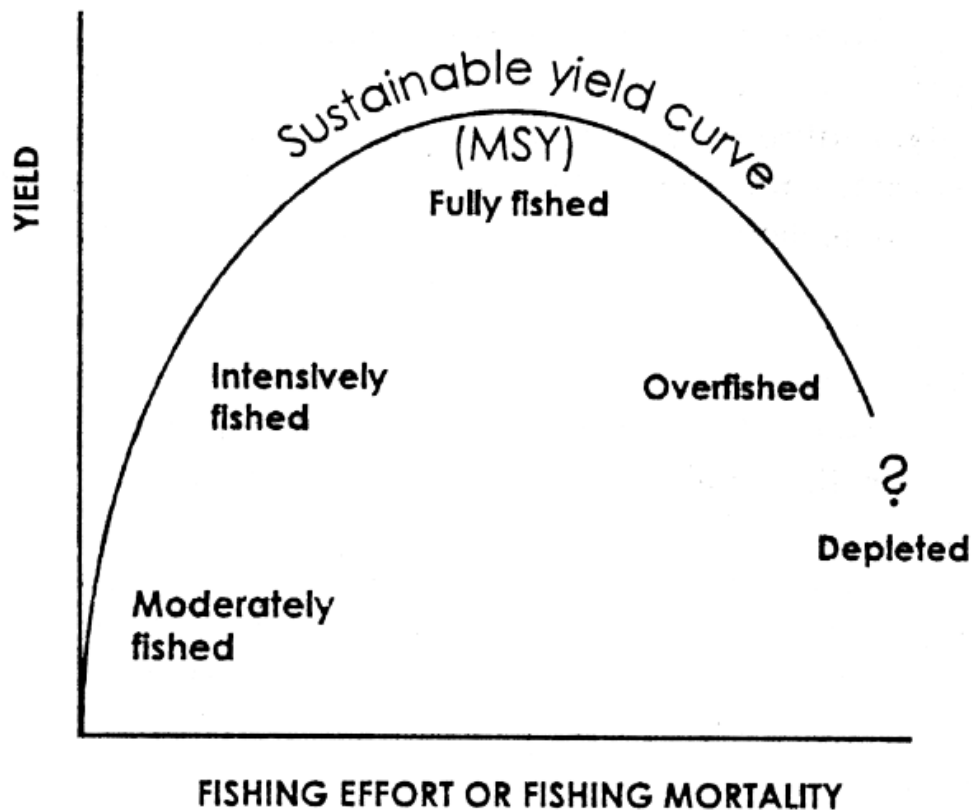


Figure 5 The five stages of fishing in relation to catch and effort where MSY is the middle stage. Before MSY is reached, fishing pressure can be considered moderate or intensive, both of which support long-term sustainability. Once fishing effort results in MSY being exceeded, the population becomes overfished or severely depleted leading to a fishery which is biologically and potentially economically unsustainable (Garcia, 1996). Reproduced with permission.

The difficulty of estimating MSY makes it a potential tool for exploratory data analysis, but limits its use much beyond that. The natural advancement is to progress to analysis of catch per unit effort, which takes catch and effort into account with the inclusion of other variables to address some of the noise which causes issues for MSY estimates.

Catch Per Unit Effort (CPUE)

The most basic and informative data in fisheries are time series of catch and effort from which catch per unit effort (CPUE) can be calculated (Teschfichael, Pitcher, & Pauly, 2014). CPUE is a common measure in fisheries science as it can be easily calculated from the compulsory administrative data collected by commercial fishers. CPUE is often considered a proxy for abundance, but there is considerable debate amongst fisheries scientists about whether CPUE is a suitable measure (Pauly, Hilborn, & Branch, 2013).

The variable *catch*, which is the weight of fish caught in an event, measured in kilograms, is considered to be a function of fishing effort and population abundance. The equation which links these components is given as follows:

$$C = qEB$$

where q is a catchability coefficient corresponding to the ability of fishermen to catch the target species, E corresponds to fishing effort which can be measured in a number of different ways and B corresponds to the relative biomass or abundance of the target species. In order to create the CPUE measure, catch is divided by effort as follows:

$$CPUE = \frac{C}{E} = qB.$$

Thus, CPUE is a relative measure of abundance defined as the product of the catchability coefficient and the species biomass (Maunder & Punt, 2004).

The units of CPUE are dependent on the choice of effort variable and are frequently made up of fishing gear or temporal variables. Some examples of CPUE units include kg per hour, kg per trawl, kg per day, kg per kilometre towed and fish per hook (Dunn, Harley, Doonan, & Bull, 2000).

Changes in CPUE through time can provide insight into changes in the abundance of the target species. A decreasing measure of CPUE is often an indicator of overexploitation, whereas constant CPUE is often an indication that the level of fishing is sustainable. The usefulness of CPUE as a measure of abundance and sustainability relies on how much is known about the catchability coefficient for the stock of interest. The simplest interpretation of the catchability coefficient is that it is constant over time and under different fishing conditions. In practice the assumption of a constant catchability coefficient is rarely met, making the proportionality relationship between CPUE and abundance a poor representation of reality.

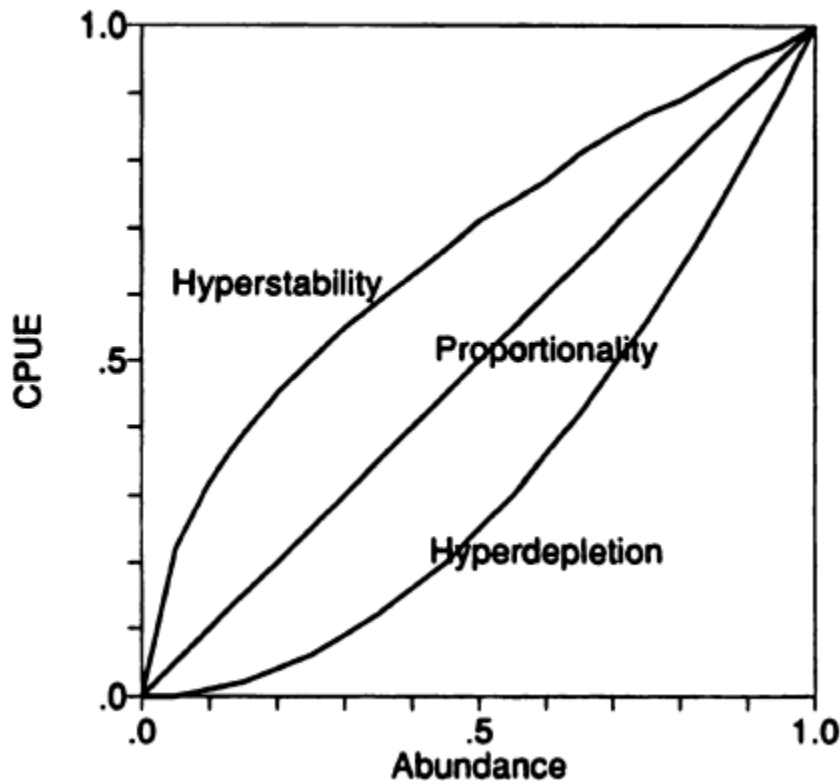


Figure 6 The three possible relationships between CPUE and abundance are hyperstability, proportionality and hyperdepletion. Proportionality is the simplest of these, but possibly the most unrealistic in practice. The choice between hyperstability, a lagged decrease in CPUE in relation to decreases in abundance, and hyperdepletion, an accelerated decrease in CPUE as abundance decreases, is dependent on factors which include the species and fishing equipment used (Hilborn & Walters, 1992). Reproduced with permission.

Two alternative relationships between CPUE and abundance are hyperstability and hyperdepletion. In the hyperstable case, CPUE remains high despite a decrease in abundance. Hyperstable relationships are expected to occur when fishing effort is focused on areas where the fish stock is abundant and is often typical of schooling species and those that form spawning aggregations. Hyperdepletion occurs when CPUE decreases faster than the decrease in abundance. This can occur when subsets of the stock are less vulnerable to fishing pressure, keeping overall abundance high. These two cases correspond to changes in the catchability coefficient over time. Pinpointing the causes and timings of these changes in q is difficult which means that estimates of abundance based on CPUE often come with considerable uncertainty. Figure 6, taken from Hilborn and Walters (1992), highlights the three possible relationships between CPUE and abundance.

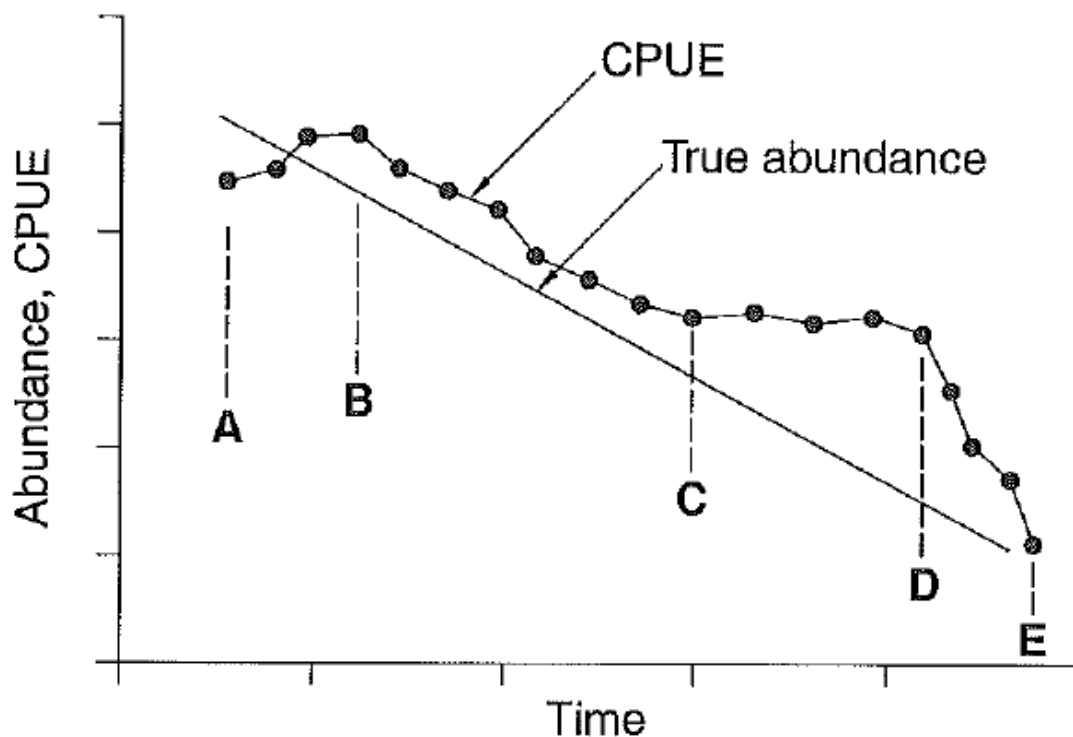


Figure 7 An example of the possible disjoint between CPUE and abundance. The CPUE in section B-C loosely follows the abundance but the introduction of more efficient fishing gear, the discovery of a new spatial aggregation of the target species, or similar leads to a period of hyperstability in section C-D where abundance is steadily decreasing while CPUE remains steady. This disjoint is corrected in section D-E, but this alignment may not come in time to prevent the stock being overexploited (King, 2007). Reproduced with permission.

Hyperstability is a double edged sword. For fishermen, hyperstability is desirable as the CPUE and their profit remain high despite decreases in abundance. For fisheries scientists, hyperstability is a cause for concern as there is potential for the stock to collapse without any warning. Some of the most well-known fisheries collapses, including North Sea herring and Peruvian anchoveta, have been attributed to hyperstability (Hilborn & Walters, 1992). Thus, while CPUE can be a useful measure, without alternative abundance measures for comparison, it should be used with caution.

Figure 7, taken from King (2007) is a hypothetical example which highlights potential discrepancies between CPUE and abundance. The time series begins with a previously unexploited stock and explores the relationship between CPUE and abundance throughout the different phases of the fishery.

The section from A – B corresponds to increasing q as the fishermen progress through the exploratory phase of the fishery and become familiar with the stock. From B – C, the fishery is at its most efficient and CPUE decreases as the stock starts to become depleted. The phase from C – D, is the one of most concern to fisheries scientists. As the main stock starts to get depleted, fishermen target the far

reaches of the stock which results in a stable CPUE for a period of time. While the CPUE stays high, the abundance continues to decrease. The phase from D – E is a correction of the previous phase where the CPUE re-aligns itself to the abundance index with a relatively sharp decrease.

The key feature to note from Figure 7 is phase C – D. Often periods where the CPUE is stable are considered to correspond to a sustainable level of fishing. The opposite is true of this section. Inspection of spatial distribution of catches and fishing efficiency before and during this phase would likely reveal a change in behaviour and sound warning bells.

Daniel Pauly and Ray Hilborn, two of the most well-known fisheries scientists, have made a name for themselves with their strong and opposing views on fisheries science, especially their views on CPUE. In February 2013, *Nature* published a debate on the suitability of CPUE as a measure of abundance (Pauly, Hilborn, & Branch, 2013). Pauly insists that CPUE data is a crucial signal as catch weight is often the only data available from many of the world's fisheries. Hilborn and his colleague, Branch, claim that CPUE is misleading as there are numerous processes that affect abundance and catch is only one of them.

There are many variables in fisheries that are likely to experience significant levels of variation. Fishing technology and equipment often become more efficient allowing fishermen to locate and catch larger quantities of fish at a faster rate. The Ministry for Primary Industries updates the total allowable commercial catch (TACC) levels to reflect abundance estimates and the commercial use and value of bycatch species may change preferred fishing locations or methods. All of these factors including a variety of others are likely to affect the CPUE. As a result of this variation, it is important to standardise the CPUE. This process aims to reduce the effect of any factors that are potentially influential on the CPUE index and the corresponding abundance estimates. The most common method for standardising a CPUE index is the application of generalised linear models (GLMs) (Hinton & Maunder, 2003).

Taking Stock of Orange Roughy

Orange roughy is shrouded in controversy. Depending on who you ask, it is either a sustainable stock which is currently under assessment for Marine Stewardship Council (MSC) certification (Marine Stewardship Council, 2014) with a final decision due in April 2016 (MRAG Americas, 2015) or New Zealand's most unsustainable fish choice at the bottom of Forest and Bird's 2013-2014 Best Fish Guide (Forest & Bird, 2014).

Orange roughy began life as an outcast of the restaurant menu. Originally named slimehead in 1957 (Jacquet & Pauly, 2008), the orange roughy did not sound like an appetising fish until remarketed in 1979 with its current name of orange roughy (Pauly, 2009). The name slimehead originates from the network of mucous producing canals found in their heads (Smithsonian National Museum of Natural History, 2015). Live orange roughy are bright red in colour with a blueish tinge (Clement Group, 2002). The bright red fades and becomes orange once the fish is dead (Flowers, 2015).

Orange roughy is known by a multitude of names which include red roughy, rosy soldierfish, deepsea perch and sea perch (Weeber, Thomas, & Dorey, 2010). In New Zealand they are also known by their Māori name, nihorota (Fish Species of New Zealand, 2016). Known scientifically as *Hoplostethus atlanticus*, orange roughy was first classified in 1889 by Collett after discovery in the north Atlantic (Marine Conservation Society, 2016) and belongs to the family Trachichthyidae (Department of the Environment, 2016).

High initial catches led to overestimates of orange roughy abundance. On one occasion, 54 tonnes of orange roughy were caught in a period of 20 minutes (Walrond, 2012). These abundant catches were short-lived and it soon became apparent that catch successes were a result of aggregating behaviour as opposed to abundance.

After the initial boom period for orange roughy fisheries around the world, orange roughy soon became a poster child for unsustainable fishing, especially in New Zealand and Australia (GoodFishBadFish, 2011). Over the past decade, sustainability concerns have led 10 of the USA's 20 largest seafood retailers to discontinue sale of orange roughy with some making public statements on the environmental impacts that led to their decision (Trenor, 2011). Celebrity chef Gordon Ramsay became aware of the sustainability concerns surrounding orange roughy, prompting him to discontinue his promotion of the species, via recipes and as an alternative to salmon, in favour of species with a more sustainable reputation (Braiden, 2015) (Clark P., 2013).

Orange Roughy Biology and Susceptibility to Overfishing

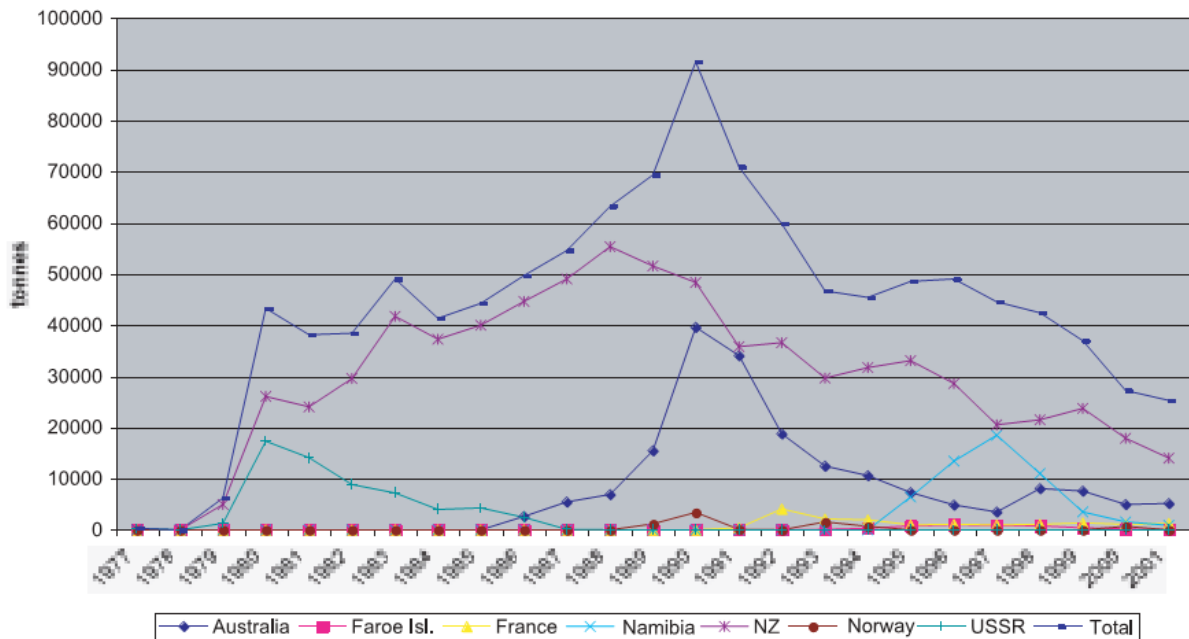
The boom and bust pattern of high initial catch, followed by substantial declines in catch resulting in fishery closures or abandonment has been seen repeatedly in orange roughy fisheries (Andrews, Tracey, & Dunn, 2009). The challenges of managing orange roughy fisheries sustainably are related to their longevity, late maturity and low fecundity. Fishers and scientists were oblivious to the unique biology of orange roughy for well-over a decade after commercial fishing of the species began.

Determining the longevity of orange roughy has been controversial (Boyer, Kirchner, McAllister, Staby, & Staalesen, 2001). Aging of orange roughy is conducted by counting the daily growth zones in the otolith (ear bone) or via radiometric analyses (Tracey & Horn, 1999). Aging of orange roughy in the early 1990s confirmed a centurion life span with a maximum confirmed age of 149 years (Fenton, Short, & Ritz, 1991), but many were not convinced due to the lack of validation of the aging methodology (Tracey & Horn, 1999). The support for a short life span for orange roughy was led by Gauldie and colleagues who disputed age estimates in the vicinity of 150+ years in favour of a much shorter life span with a maximum confirmed age of 29 years (Gauldie & Cremer, 1998) (Gauldie & Romanek, 1998).

Despite criticisms of the aging methodology, it is now widely accepted that orange roughy have a centurion life span (Doonan, Horn, & Maolagáin, 2014). The oldest estimated ages for individual fish are 187 years (Marine Conservation Society, 2016) and 194 years (Tracey & Horn, 1999). However, the standard errors around some of these age estimates are large. The 95% confidence interval for an individual with an estimated age of 40 years was 26 – 54 years (Andrews, Tracey, & Dunn, 2009).

Aging of orange roughy earns the species ninth place in the list of longest living species on the planet (Kancaid, 2015). Given their longevity, the recovery timeframe for overfished orange roughy stocks has been estimated to be 40 – 45 years (Department of the Environment, 2016) but could be as long as 65 years (Australian Fisheries Management Authority, 2014).

Orange roughy mature between 22 and 40 years of age, which exceeds the life expectancy of the majority of inshore fish species. They have an extremely low natural mortality rate and a slow growth rate. The standard length of adult orange roughy ranges from 20 – 62cm and their body mass upper limit is 7kg (Branch, 2001). These biological traits combine to make successful orange roughy management a challenge. Management of the Namibian orange roughy fishery, developed in 1994, was unsuccessful despite a precautionary approach which took the management errors of earlier orange roughy fisheries into account (Boyer, Kirchner, McAllister, Staby, & Staalesen, 2001).



Source: FAO Fishery Information, Data and Statistics Unit, 2003c

Figure 8 Reported worldwide orange roughy catch from 1977 to 2001. It is interesting to note the similarity between the time series of worldwide orange roughy catch (blue squares) with the MSY curve (Figure 5). The worldwide peak catch occurred in 1991, coinciding with the peak catch in Australia, but two years after the peak NZ catch in 1989. The NZ catch of orange roughy (purple cross) is greater than that of the other six countries combined (Lack, Short, & Willock, 2003). Reproduced with permission.

The time series of worldwide orange roughy catch mimics the MSY curve (Figure 5), suggesting that the MSY for orange roughy may have been exceeded in 1991 (Figure 8). The time series of orange roughy catch in NZ and Australia also follow the shape of the MSY curve which matches the reality that many of the orange roughy stocks in both countries have been subject to overfishing. The majority of worldwide orange roughy catch has occurred in NZ waters.

Sub-Antarctic Orange Roughy

There are many orange roughy fisheries in the Sub-Antarctic region. These fisheries have developed progressively throughout the past 25 years. The first of the Sub-Antarctic regions to experience large catches was the southeast Pukaki Rise in the 1995-96 fishing year with a total catch exceeding 3000 tonnes. After this initial boom, catches in southeast Pukaki started to drop rapidly and within a few years the fishery had entered a bust phase and fishing effort became focused elsewhere, including the northeast of the Pukaki Rise which includes the Priceless stock. This area was the dominant region for Sub-Antarctic orange roughy catch throughout the 2000s. Catches in northeast Pukaki and Priceless are focused at the start of the fishing year which runs from 1 October to 30 September. During the period from 2005 to 2010, the area catch limit of 500 tonnes was reached (Ministry for Primary Industries, 2010).

Year	Event
1990	Beginning of exploratory fishing for orange roughy in the Sub- Antarctic
1992	TACC for ORH3B is divided into voluntary catch limits managed via a voluntary industry arrangement
1995	Southeast Pukaki fishery established
1999	SE Pukaki catch decreased to less than 200 tonnes
2001	NE Pukaki and Priceless fisheries developed and Sub-Antarctic catch limit reduced to 1300 tonnes
2006	Sub-Antarctic catch limit increased to 1850 tonnes
2010	Sub-Antarctic catch limit decreased to 500 tonnes

Table 2 Timeline of development of the orange roughy fishery in the Sub-Antarctic from 1990 to 2010
(Ministry for Primary Industries, 2010)

Decreases in the landings and TACC for ORH3B are representative of the boom and bust nature of orange roughy fisheries (Figure 9).

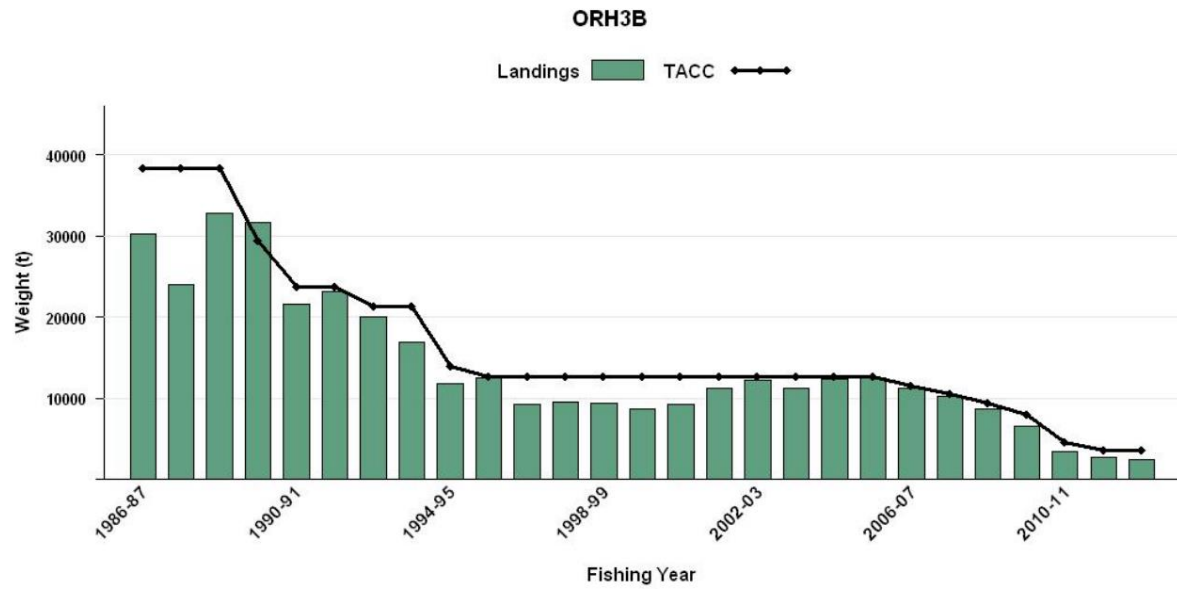


Figure 9 Total landings and TACC for ORH3B QMA from 1986 – 2013. Large landings (landed catch) occurred in the early phases of the ORH3B fishery. Both landings and the TACC decreased until a period of stability from 1995 – 2006, before another period of decrease. Catches in the Priceless region are included in the ORH3B TACC (Ministry for Primary Industries, 2014). Reproduced with permission.

Methodology – Fish & Statistics

Data Overview

This project utilises commercial fisheries data from MPI for all reported orange roughy and oreo catches in New Zealand since the beginning of reporting in the fishery. This dataset forms part of the New Zealand Catch Effort database *fish_ce*. This catch data and the associated variables were recorded by fishers on official MPI forms which include the Trawl Catch Effort Processing Return (TCP), High Seas Trawl Catch Effort Return (HTC) and Catch Effort Landing Return (CEL) forms. The combination of effort, environmental and administration variables recorded for each fishing event depends on the vessel type and species being targeted. Orange roughy and oreo are deep-water species usually caught by bottom trawling so the variables in the commercial catch data reflect this.

There are four main commercial catch datasets. Each of these consists of a different set of variables but can be combined using unique ID variables. The most relevant dataset for this project is the fishing event dataset. The fishing event dataset is made up of spatial and temporal data relating to catch and fishing effort for each event. The catch variable in the fishing event dataset is the weight of the entire catch in kilograms. The estimated sub-catch dataset contains estimated green weights for each of the species included in the catch. The processing event dataset contains processing details for catch that is processed out at sea. The landings dataset records landing information either as raw catch or processed catch. These four datasets can be linked using *event key* or *trip*, which are unique identifier variables.

The two most relevant of these commercial datasets to stock assessment are the fishing event and estimated catch datasets. Merging the fishing event and estimated catch datasets using the *event key* variable enables use of the spatiotemporal, effort and environmental variables from the fishing event data while also utilising the more accurate catch and species variables from the estimated catch data.

Variables

The combined dataset has many variables which are superfluous for analysing orange roughy and oreo catch. As a result, many of these variables were removed to reduce the size of the dataset. The meaning of some of the variables is self-explanatory but for others a definition or units are required to fully understand the variable. Interpretation of the variables in the database was aided by referral to two database documentation reports (Mackay, 2005) (Research Data and Reporting Group, 2010). The most relevant variables from the fishing event and estimated catch datasets are summarised to facilitate discussion of models and analysis in further sections (Table 3).

Variable Name	Description and variable format
event_key	Unique fishing event number
start_datetime	Start fishing date and time using 12 hour clock In format: dd/mm/yyyy hh:mm:ss.000 a.m./p.m.
end_datetime	End fishing date and time using 12 hour clock In format: dd/mm/yyyy hh:mm:ss.000 a.m./p.m.
target_species	Three letter target species code e.g. ORH
fishing_duration	Duration of fishing event measured to 3dp in hours in format 0.000
catch_weight	Estimate of the total weight of the catch for the fishing event measured in kg
effort_depth	Depth of effort in m measured as ground rope depth.
effort_height	Effort height measured as headline height in m where headline corresponds to the rope which runs along the upper mouth of the net
effort_width	Effort width is measured as the wing spread of the trawl net in m.
effort_speed	Estimated speed of trawl
start_latitude	Start latitude for event in decimal degrees format
start_longitude	Start longitude for event in decimal degrees format
end_latitude	End latitude for event in decimal degrees format
end_longitude	End longitude for event in decimal degrees format
bottom_depth	Depth of sea bottom in m
display_fishyear	Formatted fishing year e.g. 1 Oct 1996 to 30 Sep 1997 = 1996/97
trip	A system generated number allocated to each of the events that took place for one vessel between its trip start and end dates
vessel_key	Number generated by MPI to identify the vessel fishing
surface_temp	Sea surface temperature in °C
bottom_temp	Sea bottom temperature in °C
species_code (est_catch)	Three letter code identifying the species caught
catch (est_catch)	Estimated catch of the species in kg

Table 3 Summary of useful commercial catch variables for the Priceless region used for exploratory analysis and modelling. These variables are from the fishing event and estimated catch datasets.

Some of the variables in the table above are more useful in a different format. The `start_datetime` and `end_datetime` variables play an important part in the investigation of catch and effort trends over time. These variables are considerably more useful when they are broken down into their individual components. The resolution of day and time of day are more precise than what is required to analyse trends in catch and effort so can be removed from the dataset. The temporal levels of resolution which will be of most use are year and month. Since the majority of events start and end on the same day, `end_datetime` is removed from the dataset and year and month variables are created from `start_datetime`. For modelling purposes it is interesting to treat year and month as both factors (categorical) and numeric variables thus temporal variables of both types were created. The variables `year_F` and `month_F` indicate the variables are treated as factors while `year_N` and `month_N` are used to denote a numeric representation. A selection of other temporal variables were created including `date1960` which is a continuous variable corresponding to the number of days since 1 January 1960 and `YrMonth` which is a categorical variable consisting of year and month concatenated into a single variable. These two variables are a less natural representation of time than year and month so less focus was placed on these variables.

In the Priceless region the two species which are targeted by fishers are orange roughy and oreo. Due to the nature of deep-sea trawling, non-target species are often caught simultaneously. These non-target species are referred to as bycatch and the majority of them are not considered to be commercially significant. The estimated catch dataset contains catch weight estimates for all species caught but orange roughy and oreo are the only species of interest in this analysis. Thus, in order to simplify the species variable, a variable called `three_species` is created which classifies each species into one of three categories, orange roughy, oreo or other. The other category groups together species caught as bycatch. Thirty-nine different species were caught in the Priceless region during the period included in the dataset. This includes orange roughy, four species of oreo which include smooth oreo (*Pseudocyttus maculatus*, SSO), black oreo (*Allocyttus niger*, BOE), warty oreo (*Allocyttus verucosus*, WOE) and spiky oreo (*Neocyttus rhomboidalis*, SOR), and 34 bycatch species of which rattails are the most common. The smooth oreo and black oreo are the most commonly caught of the oreo species. When fishers target oreo they are targeting smooth or black oreo. Sometimes they will specify the species on the form but the collective classification of oreo is commonly used instead. Over half of the fishing events in the Priceless region are targeting orange roughy.

One of the most important created variables is CPUE where CPUE is Catch Per Unit Effort. This is calculated by scaling catch by the duration of the fishing event. This variable is used as a response variable to model the relationship between catch and effort over time.

Southern Oscillation Index (SOI)

The Southern Oscillation Index (SOI) is a standardised index based on observed differences in the sea level pressure (SLP) between Tahiti, an island in the South Pacific, and Darwin in Australia's Northern Territory. The SOI measures large scale fluctuations in air pressure which occur between the western and eastern tropical Pacific Ocean. The smoothed time series of the SOI closely follows the time series of ocean temperatures across the eastern tropical Pacific (NOAA, 2016) .

El Niño and La Niña refer to the two extremes of the SOI. Prolonged periods of negative SOI values correspond to warmer water temperatures across the eastern tropical Pacific and are referred to as El Niño. Conversely, La Niña is categorised by prolonged periods of positive SOI values which correspond to cooler water temperatures. Values of the SOI close to zero are classified as neutral.

El Niño episodes typically occur about 3 to 7 years apart. An El Niño episode typically begins around April or May with conditions persisting for around a year (Wratt, Basher, Mullan, & Renwick, 2013).

The SOI is calculated as follows:

$$SOI = \frac{\text{Standardised Tahiti SLP} - \text{Standardised Darwin SLP}}{\text{Monthly standard deviation}}.$$

An expansion of the standardised terms for both locations in the SOI formula is given as follows:

$$\text{Standardised Tahiti} = \frac{\text{Actual Tahiti SLP} - \text{Mean Tahiti SLP}}{\text{Standard Deviation of Tahiti SLP}}$$

$$\text{Standardised Darwin} = \frac{\text{Actual Darwin SLP} - \text{Mean Darwin SLP}}{\text{Standard Deviation of Darwin SLP}}$$

$$\text{Monthly standard deviation} = \sqrt{\frac{\sum (\text{Standardised Tahiti} - \text{Standardised Darwin})^2}{N}}$$

where N is the number of months that the index is summed over (NOAA, 2016).

The usual range for the SOI is between -3.5 to +3.5. There is also a convention that involves multiplying the SOI formula by 10 which results in whole number values ranging from -35 to +35 (Commonwealth of Australia, Bureau of Meteorology, 2016).

The SOI is usually computed monthly, but can also be computed at different temporal scales including daily, weekly and yearly. Daily and weekly SOI are less reliable as they are more susceptible to fluctuations from weather patterns. As a result, climate analyses focus on the monthly and yearly indices (Commonwealth of Australia, Bureau of Meteorology, 2016).

Due to New Zealand's geographical location, El Niño conditions are weaker than those experienced throughout Australia and other countries located in the more tropical regions of the Pacific Ocean. Despite this, El Niño still has an important influence on the New Zealand climate which warrants management and planning preceding and during El Niño episodes. The waters surrounding the country are prone to cooler temperatures during an El Niño episode and warmer conditions during a La Niña episode. La Niña conditions have a weaker effect on the New Zealand climate than that of El Niño (Wratt, Basher, Mullan, & Renwick, 2013).

A time series of the monthly SOI with the corresponding El Niño / La Niña / neutral classification spanning the years 1998-2012 was extracted from NOAA (Climate Prediction Center Internet Team, 2015). This dataset was merged with the Priceless dataset in R to allow for modelling of CPUE and sea surface temperature with the SOI.

Priceless Region Boundary

The raw dataset consists of all orange roughy and oreo catch within New Zealand's EEZ. The focus region for this study is the Priceless region which is located on the northern side of the Pukaki Rise within the region known as North Pukaki. This region lies south-east of the base of the South Island within Sub-Antarctic waters.

The Priceless region is a rectangular box defined by the following latitude and longitude coordinates:

Degrees, minutes and seconds (DDD, MM, SS)	Decimal degrees (DDD.DDDD)
48°05.5' S, 174°42' E	(-48.0917, 174.7000)
48°05.5' S, 175°13' E	(-48.0917, 175.2167)
48°26.5' S, 175°13' E	(-48.4417, 175.2167)
48°26.5' S, 174°42' E	(-48.4417, 174.7000)

Table 4 Definition of the Priceless region with coordinates in degrees, minutes and seconds (DDD, MM, SS) and Decimal Degrees (DDD.DDDD) (Anderson & Dunn, 2012) (CSGNetwork, 2016).

Latitude gives an indication of location with respect to the equator. The latitude at the equator is zero while the latitude at the North Pole is 90° north and the latitude of the South Pole is 90° south or -90°. The metric distance between each degree of latitude decreases as the location becomes closer to either of the poles.

Lines of longitude extend from the North to South Pole with a longitude of zero occurring at the Prime Meridian which runs through Greenwich, England. Movements to the east of the Prime Meridian lead to positive longitude values in the range of 0 to 180 and movements to the west correspond to negative values in the range 0 to -180 (CSGNetwork, 2016).

The degrees, minutes and seconds (DDD, MM, SS) representation is the conventional method for recording latitude and longitude coordinates. The start and end trawl locations, as recorded in the dataset, are in decimal degrees format where the minutes and seconds components are displayed as a decimal. A GPS conversion tool was used to convert coordinates from degrees, minutes and seconds to decimal degrees (CSGNetwork, 2016). These new coordinates were then used to create subsets of the data corresponding to catch in the Priceless, North Pukaki and Spawning Box regions. Each subset consists of trawls where the start and/or end coordinates fall within the region boundary.

Degrees, minutes and seconds (DDD, MM, SS)	Decimal degrees (DDD.DDDD)
47°30' S, 176°18' W	(-47.5000, 176.3000)
47°30' S, 173°00' W	(-47.5000, 173.0000)
49°00' S, 173°00' W	(-49.0000, 173.0000)
49°00' S, 176°18' W	(-49.0000, 176.3000)

Table 5 Definition of the North Pukaki region, located in the Sub - Antarctic, with coordinates in degrees, minutes and seconds (DDD, MM, SS) and Decimal Degrees (DDD.DDDD). The North Pukaki region is a rectangle defined by the coordinates in the table excluding the area defined as the Priceless region (Anderson & Dunn, 2012) (CSGNetwork, 2016)

Degrees, minutes and seconds (DDD, MM, SS)	Decimal degrees (DDD.DDDD)
42°10' S, 178°00' W	(-42. 1667, 178.0000)
42°10' S, 175°00' W	(-42. 1667, 175.0000)
44°00' S, 175°00' W	(-44.0000, 175.0000)
44°00' S, 178°00' W	(-44.0000, 178.0000)

Table 6 Definition of the Spawning Box region, located within the East and South Chatham Rise, with coordinates in degrees, minutes and seconds (DDD, MM, SS) and Decimal Degrees (DDD.DDDD). The Spawning Box is one of the key orange roughy stocks located in the East and South Chatham Rise (Ministry of Fisheries, 2008) (CSGNetwork, 2016).

The Priceless, North Pukaki and Spawning Box stocks are all located in the orange roughy Quota Management Area (QMA) known as ORH3B (Figure 10).

The NZ EEZ dataset refers to all orange roughy and oreo catches recorded in the MPI database. The dataset includes catches from October 1989 to December 2014 which corresponds to a period of just over 24 years. Catches continue to the present day with the focus being on orange roughy fisheries in the Chatham Rise (ORH3B) and Challenger Plateau (ORH7A). This dataset thus consists of orange roughy fisheries that are considered to be both healthy and overfished and the exploratory analysis of this data allows for a crude analysis of the overall health of fished orange roughy stocks in New Zealand.

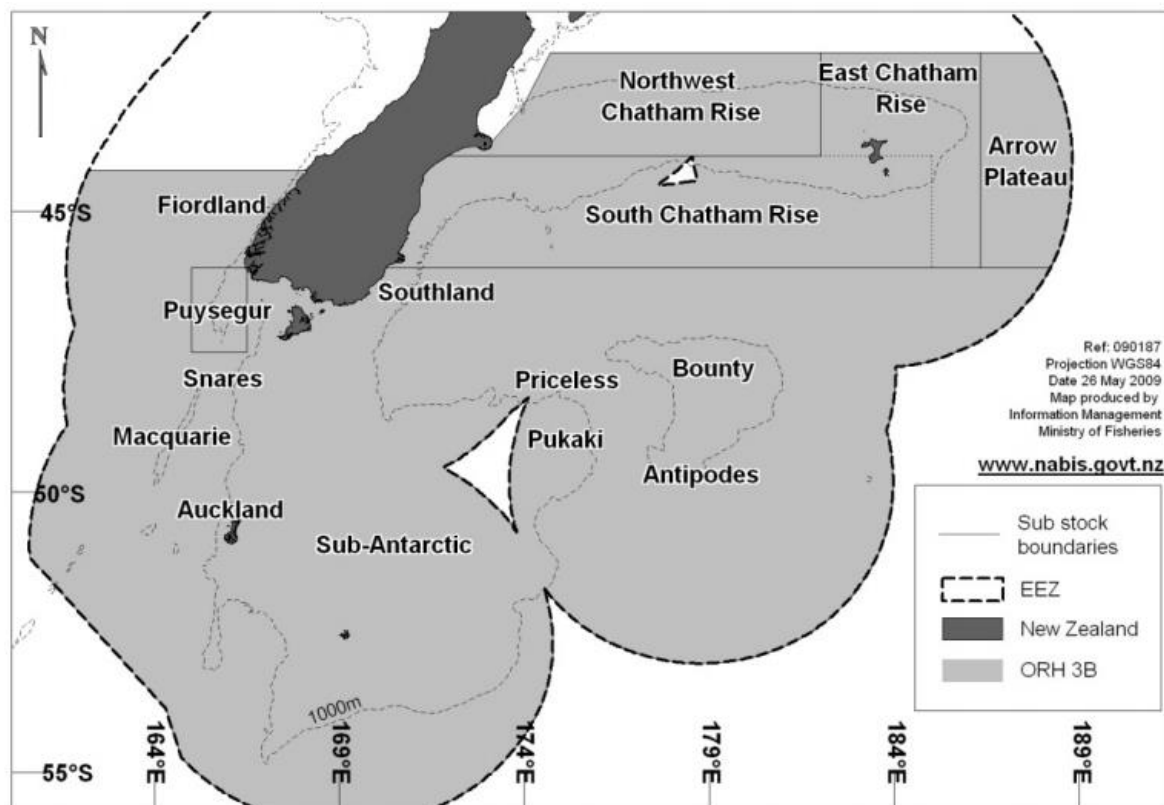


Figure 10 Map of the ORH3B QMA consisting of the Chatham Rise and Sub-Antarctic stocks. The Priceless, North Pukaki and Spawning Box stocks all belong to this region. The map includes bathymetric contours and the EEZ boundary (Ministry for Primary Industries, 2014) Reproduced with permission.

The map of catch locations was created in R using ‘mapdata’ and plotting the start latitude and longitude for each event in the dataset (Figure 11). Measurement errors in the latitude and longitude coordinates have resulted in some of the catch locations occurring on the mainland. These measurement errors can result in events being misallocated between stocks or in the case of those events recorded on the mainland, not allocated to a stock at all.

When plotting trawl locations (Figure 24), any start or end locations which have been incorrectly measured need to be identified and dealt with appropriately. These measurement errors can be identified based on whether the distances travelled are realistic given the fishing duration and vessel speed. Once location outliers have been identified, a new coordinate can be imputed using the average trawl distance for events in the region. There were five location measurement errors detected for the Priceless data. Because there were so few errors, imputation was considered superfluous and these data points were ignored in the creation of Figure 24 and other analyses involving location.

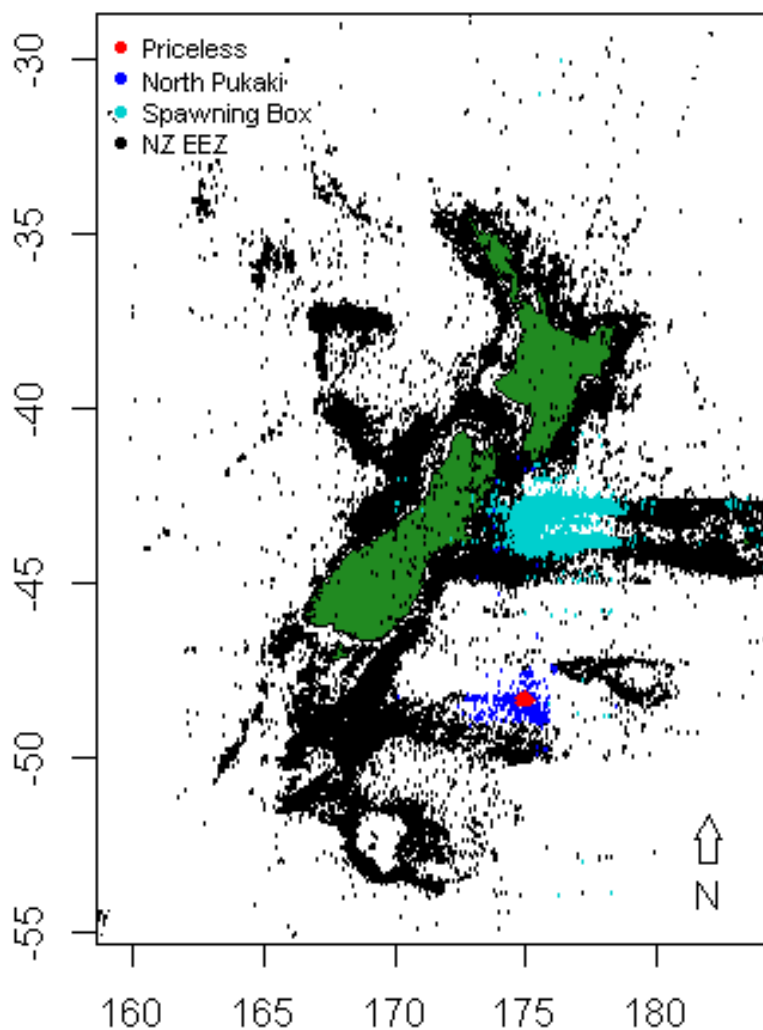


Figure 11 Map showing the start locations of fishing events involving orange roughy and oreo catch in the NZ EEZ. Locations of catches in the Priceless, North Pukaki and Spawning Box regions are highlighted for reference. The catch locations plotted on the NZ mainland are the result of measurement errors.

In the Priceless region all of the data has been collected on Trawl Catch Effort Processing Return (TCP) forms. There are 1239 events in the Priceless region where orange roughy was included in the catch. The primary fishing method for all of these events was bottom trawling.

Defining CPUE for Priceless Orange Roughy

CPUE analysis for New Zealand fish stocks often uses measures of effort such as fishing day, distance towed and hooks set, depending on the species of interest. This leads to CPUE measures in the following units: kg per day, kg per kilometre towed, and fish per hook (Dunn, Harley, Doonan, & Bull, 2000). Orange roughy is caught by bottom trawling so measures involving hooks are irrelevant. Given that the data is collected for individual fishing events, using fishing day as the effort measure seems unnecessarily coarse. The effort measure used here is fishing duration which is measured in hours. Thus, the units of CPUE defined using fishing duration become kg per hour. The duration of a fishing event is highly correlated with distance towed, so it is expected that the time series of CPUE will perform similarly for both of these effort measures.

Due to the nature of commercial catch data and the small quantities of bycatch included in the data, the CPUE effort measure is dominated by relatively small catches. In good conditions, catches can be significantly higher than average, which leads to a large range. The overall skewness of the CPUE data makes it harder to identify patterns and trends. It is therefore beneficial to log transform the CPUE data.

There are two common methods for dealing with zero catch when analysing CPUE data. The first of these is to stratify CPUE into groups but this method results in a loss of information. The second method commonly used is to add a constant to the data so that the CPUE is always greater than zero. A disadvantage of using this method is that there is a possibility of introducing bias (Hinton & Maunder, 2003). CPUE is modelled here at the individual event level rather than taking the average CPUE by month or year. As a result, the constant 0.01 is added to the CPUE variable before taking the log to avoid issues with low and zero catches. An event with a fishing duration of zero is removed from the dataset to avoid an undefined CPUE.

Linear models (LMs) and generalised linear models (GLMs) are commonly used in fisheries science when analysing CPUE. Their purpose is to identify the factors which contribute to variation in the CPUE index so that their effect can be removed from the relative abundance index.

There are many explanatory variables which could have an effect on CPUE. In fisheries with well-established stock assessment programs, it is often known which variables are most likely to have an effect on the CPUE index. This allows for fitting of a selection of LMs with the known explanatory variables. In lesser studied fish stocks, the explanatory variables may not be known. This leads to a preliminary stage of modelling which is completed to determine the variables of interest for further modelling and prediction.

Modelling CPUE using Linear Models (LMs)

Linear regression is used to study dependence relationships between a set of explanatory variables and a response variable. LMs are fitted to data using the method of least squares which provides estimates of both the regression coefficients and their standard errors. The method of least squares calculates the line of best fit between the response and explanatory variables by minimising the sum of squares of the residuals. The LM algorithm in R solves for the best fit using QR decomposition. Once the line of best fit has been found, it can be used for explanation and prediction.

LMs consist of two main components:

1. A linear predictor which is a linear function of the explanatory variables. The linear predictor is written in the following form:

$$E(Y_i) = \alpha + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik}.$$

Here α represents the intercept, the β_i are coefficients and the X_{ik} are the explanatory variables included in the model.

2. A random component which specifies the conditional distribution of the i th response variable Y_i out of n independently sampled observations, given the values of the explanatory variables included in the model.

Simple LMs include one explanatory variable, while multiple LMs include more than one explanatory variable and may also include interactions between variables.

GLMs are a generalisation of ordinary linear regression which allow for non-normal error distributions for the response variable. When the errors are normally distributed, the link function is the identity, making the normal GLM identical to the standard LM. Standard LMs are the most common type of regression used in practice. Linear models are used to model the CPUE data as the assumption of normally distributed residuals is met.

In R, linear models can be inputted using the following notation: `linear_model <- lm(Y ~ X)` where Y is the response variable and X is a linear function of the explanatory variables. The R syntax for entering individual variables and interactions to the function of explanatory variables is summarised in Table 7.

R Syntax	Meaning
+ X	include the variable X
- X	do not include the variable X
X:Z	include the interaction between the variables X and Z
X*Z	include X and Z and the interactions between them
X Z	conditional statement to include X given Z
(X+Z+W)^3	include these three variables and all their interactions up to three way
I(X*Z)	include a new variable consisting of the quantity in brackets

Table 7 R syntax for including explanatory variables in a linear model.

LM assumptions and diagnostics

There are a set of assumptions which accompany LMs. These assumptions can be expressed and validated in a variety of ways. The six assumptions which follow all relate to the residuals.

1. The mean of the residuals is zero.
2. The residuals are homoscedastic (i.e. the variance of the residuals is constant).
3. The residuals are normally distributed.
4. The residuals are independent.
5. The residuals show no evidence of departure from linearity.
6. The residuals indicate the independent variables have a fixed distribution.

These assumptions are summarised from Verran & Ferketich (1987). The first four of these are sometimes referred to as explicit assumptions, with the final two often considered to be implicit as they are implied by the specification of the model. The explicit assumptions can be validated by plotting model diagnostics. Two of the key diagnostic plots for LMs are displayed (Figure 12). The plot on the left checks for homoscedasticity. The residuals are plotted against the fitted values. Ideally the plotted points will appear randomly scattered with no patterns present. Obvious patterns in this plot are a sign that the assumption of homoscedasticity is violated. The plot on the right is a normal quantile-quantile plot (Q-Q plot) which is used to check the assumption of normality of the residuals. Ideally the points will all fall in a line. Large deviations from this line indicate that the normality assumption has been violated. The quantiles in the Q-Q plot deviate from the straight line in the lower and upper tails which is not ideal (Figure 12). However, LMs are reasonably robust to violations of the normality assumption. If this assumption is not met, the p-values will be unreliable but estimates of the coefficients will still be unbiased.

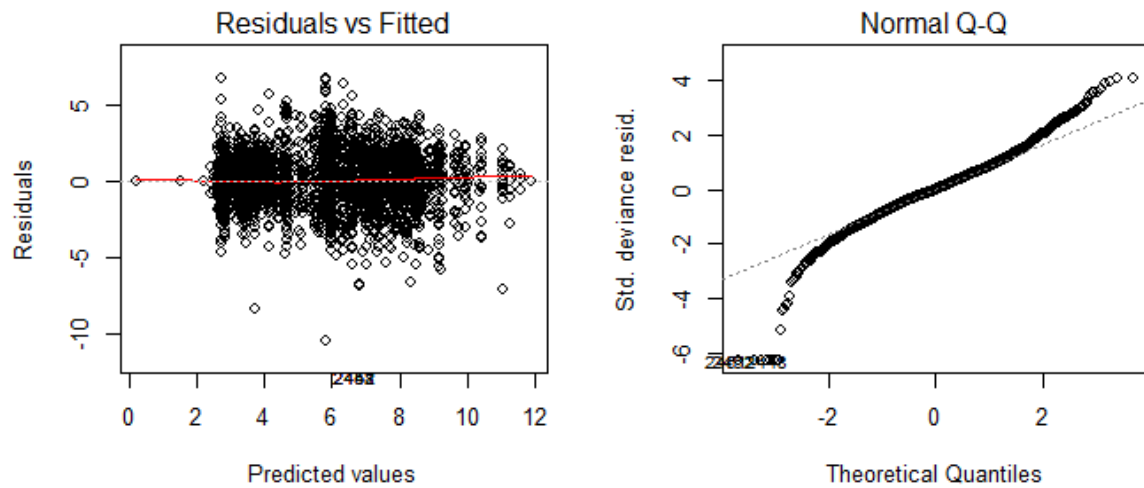


Figure 12 Two diagnostic plots for testing the assumptions of linear models. On the left, the homoscedasticity assumption is tested by plotting the residuals against the fitted values. A good result for this plot is random scattering around the line at zero. The diagnostic plot on the right tests for normality of the residuals. The points will fall in an approximately straight line if this assumption holds.

As well as diagnostic plots, there are a variety of tests that can be run to check the assumptions of LMs.

Modelling CPUE using Temporal Variables

The LMs fitted for CPUE include many explanatory variables, all of which place different assumptions on the behaviour of the response variable. It is important to be aware of these assumptions and their ability to represent reality. A selection of LMs were run based on all possible combinations of year and month variables to determine which of the assumptions is the most valid in the given context.

Year and month can be fitted as categorical or continuous variables (year_F, month_F, year_N and month_N). Modelling with categorical variables assumes abrupt fluctuations in the response variable for each category of the variable (Figure 13A, B). The mathematical equation corresponding to this model is:

$$E(y) = \beta_x$$

where x is a categorical variable and β_x represents the set of coefficients for each level of the categorical variable. If x includes n categories, this model involves the estimation of n coefficients.

Continuous models assume a linear trend in the response variable (Figure 13C, D) and are represented by the following LM:

$$E(y) = \alpha + \beta x$$

where x is a continuous variable, α corresponds to the intercept of the model and β to the coefficient of x . This model involves the estimation of one coefficient.

When using the continuous month variable it is possible to incorporate seasonality. This can be achieved by including harmonic terms in the LMs. Because there are twelve months per year it makes sense to model up to the fifth harmonic.

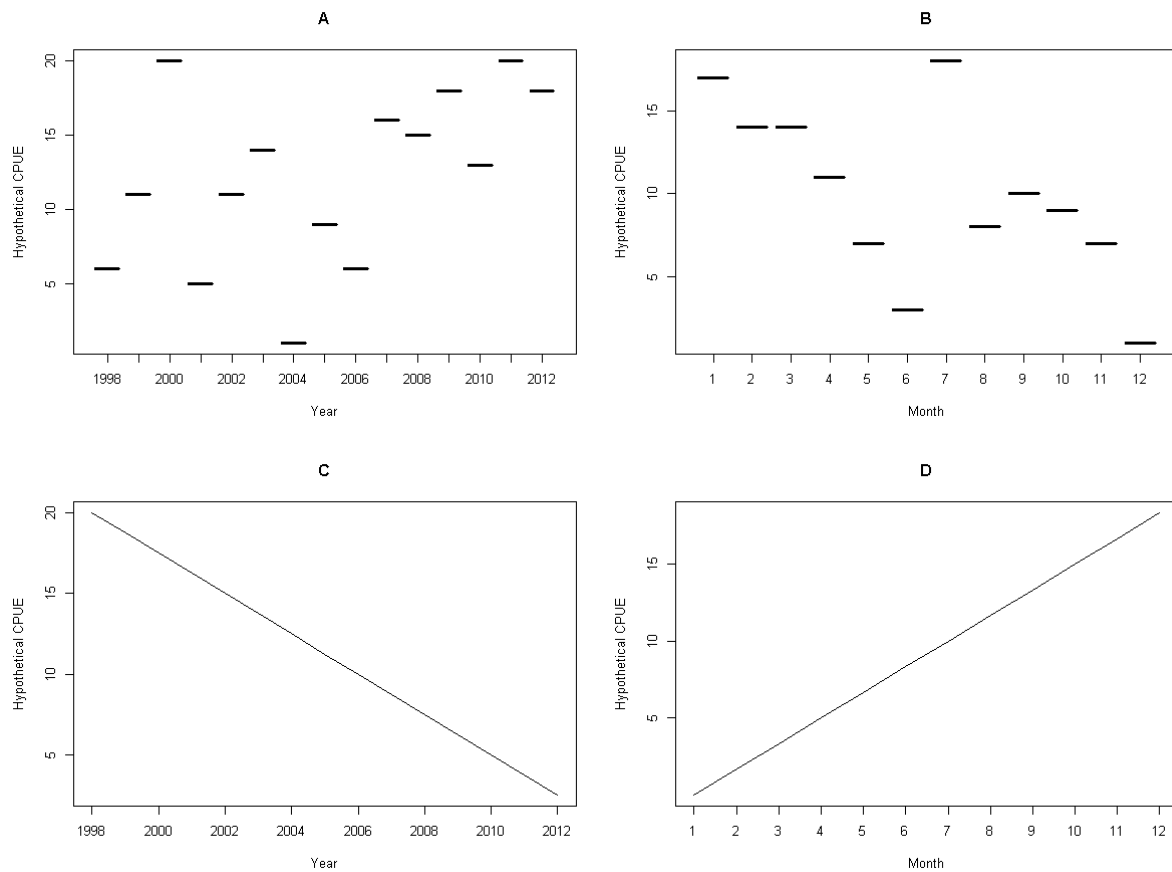


Figure 13 Simulated graphs to represent the underlying assumptions of modelling CPUE with year and month as categorical or continuous variables. Plot A demonstrates annual fluctuations in CPUE when modelled with year as a categorical variable. Plot B demonstrates monthly fluctuations in CPUE when modelled with month as a categorical variable. The fluctuations resulting from these categorical variables are unpredictable and do not follow an obvious pattern. Plot C demonstrates a decreasing linear trend in CPUE which results when year is a numeric variable. Similarly, Plot D demonstrates a monthly linear trend in CPUE.

The monthly harmonics are composed of trigonometric functions. Thus the decomposition of annual seasonal patterns is attained using a trigonometric series.

The general form for a harmonic series involves a sample of T observations y_1, \dots, y_T and a trigonometric expression of the form:

$$z_t = \sum_{j=1}^n \{\gamma_{1j} \sin(\omega_j t) + \gamma_{2j} \cos(\omega_j t)\}$$

where $\omega_j = 2\pi j/T$ denotes the frequency and γ_{ij} are the coefficients of the series where $j = 1, \dots, n = \frac{T}{2}$ (Pollock, 2011).

A model for annual seasonality could include one peak and trough or up to six peaks and troughs. The case of six peaks and troughs is equivalent to treating month as a categorical variable. Thus, the analysis to follow goes up to the fifth harmonic (Figure 14A – E).

Here, $T = 12$ corresponding to the 12 months in a year. The coefficients, γ_{ij} , are equal to one and j ranges from 1 to 5. Note that the sixth harmonic corresponds to the $n = \frac{T}{2}$ case.

The harmonic terms to model seasonality in CPUE are written as:

$$month_N_n = \sum_{j=1}^n \sin\left(\frac{2\pi j * month_N}{12}\right) + \cos\left(\frac{2\pi j * month_N}{12}\right).$$

Table 8 shows the formulas for the monthly harmonics where N_j corresponds to the j th harmonic which is made up of j peaks and troughs.

Month harmonic	Formula _j
Month_N1	$\sin(2 \pi * month_N/12) + \cos(2 \pi * month_N/12)$
Month_N2	$Month_N1 + \sin(4 \pi * month_N/12) + \cos(4 \pi * month_N/12)$
Month_N3	$Month_N2 + \sin(6 \pi * month_N/12) + \cos(6 \pi * month_N/12)$
Month_N4	$Month_N3 + \sin(8 \pi * month_N/12) + \cos(8 \pi * month_N/12)$
Month_N5	$Month_N4 + \sin(10 \pi * month_N/12) + \cos(10 \pi * month_N/12)$

Table 8 Formulas for the monthly harmonic variables

An LM incorporating monthly harmonics is formulated as follows:

$$E(y) = \alpha + \sum_{j=1}^n \left\{ \beta_{1j} \sin\left(\frac{2\pi j x}{12}\right) + \beta_{2j} \cos\left(\frac{2\pi j x}{12}\right) \right\}$$

where the number of coefficients, β , to be estimated is $2n$.

To include interactions involving the harmonic month terms, brackets are placed around the trigonometric sum so that interactions are fitted to the full term.

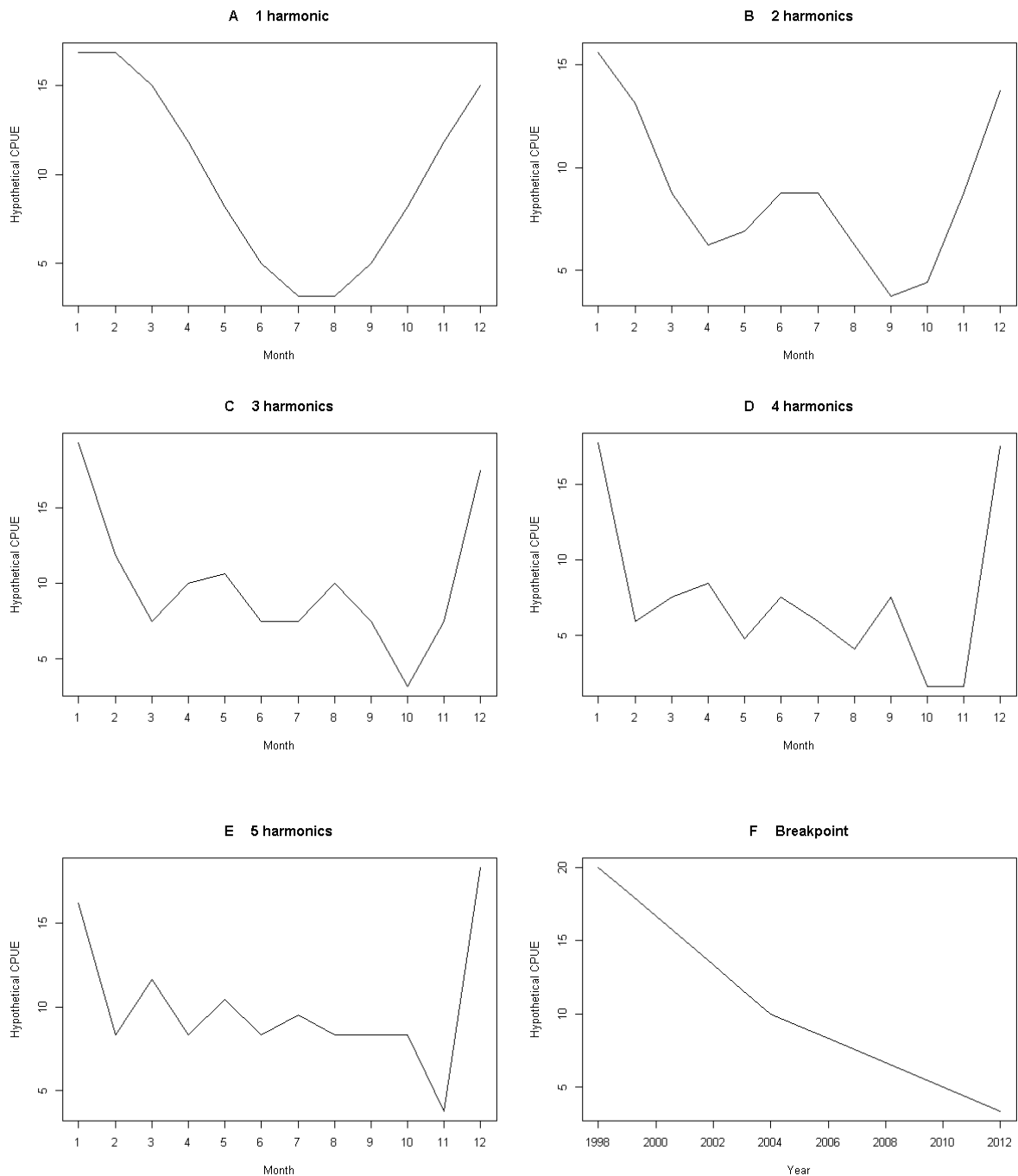


Figure 14 Simulated graphs to represent the behaviour of the monthly harmonic variables and a segmented linear trend. Plot A to E demonstrate the first five monthly harmonics which include one to five peaks and troughs within each year respectively. Plot F demonstrates the behaviour of the breakpoint model. The breakpoint model occurs when there is an annual linear trend in CPUE (Figure 13C) where the slope of this trend is different either side of the breakpoint. The breakpoint model is discussed presently.

Segmented Linear Models and the Identification of Breakpoints

Segmented linear regression (also known as piecewise, change-of-phase, hockey stick and join point regression) is a useful extension of linear regression which takes into account the possibility of more than one linear relationship in the data and can help identify critical thresholds within the data (Brendan & Bence, 2008). These linear relationships are separated by a sudden change in directionality, known as the breakpoint.

Segmented regression models have been used in fisheries science, including modelling size and age related mortality (Maceina, 2007). Critical thresholds which require segmented regression are common in ecological processes, especially the modelling of species or habitat loss or management regimes through time (Toms & Lesperance, 2003). Analysis on reported catch in world fisheries and reconstructed catch rates, which include estimates of illegal and unreported catch, were modelled using segmented regression and identified two separate breakpoints. These breakpoints correspond to the years where the slope of the catch trend experienced a significant change (Pauly & Zeller, 2016).

There are a variety of methods for identifying breakpoints for segmented regression models but one of the most widely used is the grid-search method (Brendan & Bence, 2008). This method involves a grid of potential candidate values for the breakpoint. A linear model is then fitted with each of the candidate breakpoints and a criterion is set to select the breakpoint which corresponds to the best model fit (Muggeo, 2008). While this method is straight forward to implement, it is not without issues. Given that the method involves fitting models for each of the breakpoint candidates, this method can become computationally expensive with large datasets and the algorithm is potentially cumbersome when the locations of multiple breakpoints are to be estimated. Estimating the breakpoint as a fixed value can also lead to issues with the standard errors of other parameters being underestimated as the uncertainty in the breakpoint is not considered in the algorithm (Muggeo, 2008).

The process of identifying and fitting a segmented regression can be executed using the R package 'segmented'. This involves fitting an LM (or GLM) to the data and then specifying the variable with the suspected breakpoint and an estimate of the location of this breakpoint. Once the best breakpoint has been identified, the previously fitted model is updated to include the segmented relationship.

The location of the breakpoint and the slopes of the linear relationships either side of this point are the key components of interest. In order to fit a segmented regression, the variable, x , to be fitted with a breakpoint is inputted as follows:

$$E(y) = \alpha + \beta_1 x + \beta_2 (x - BP) I_{\{x \geq BP\}}$$

where I is an indicator function with an output of one if the input is true and zero otherwise. BP takes the value of any of the candidate breakpoints during the identification stage and the best breakpoint for the final model fit.

The breakpoint in annual CPUE occurs when year is defined as a numeric variable. The underlying assumption made when fitting a breakpoint with year as a numeric variable is that there is a decreasing linear trend in CPUE by year up to the year of the breakpoint and post this point the slope of the linear trend changes.

Some of the models which include the breakpoint also include interactions. Interactions with the breakpoint variable are inputted by placing brackets around the breakpoint components.

Model comparisons to identify the best model or best breakpoint can be done using AIC, the use of which is discussed subsequently.

Model Comparisons using Akaike's Information Criterion (AIC)

The Akaike Information Criterion (AIC) is one of many statistical measures which can be used to evaluate the relative quality of statistical models which are fitted to the same dataset. Using AIC for model comparison thus allows for the selection of the best model. The statistically best model is the one with the lowest AIC value.

The formula for AIC is

$$AIC = -2(\ln(L)) + 2k$$

where L is the maximum likelihood for the model under assessment and k is the number of parameters in the model (Akaike, 1974).

The likelihood component of the AIC equation summarises the descriptive ability of the model while the addition of $2k$ to the formula acts as a penalty for the inclusion of more parameters. When comparing a selection of candidate models, the trade-off between descriptive ability and parsimony needs to be considered.

Raw AIC values can be difficult to interpret, especially when there are many models and the AIC values consist of five significant figures as in the results section of this report. A solution to this is to transform the raw AIC values to enable comparison of the relative performance of the models (Wagenmakers & Farrell, 2004).

The first step of this process is to calculate the ΔAIC , the differences between the AICs of each model and the best model.

$$\Delta_i(AIC) = AIC_i - \min AIC.$$

Using the ΔAIC values, an estimate of the relative likelihood (L) of model i (M_i) can be calculated as follows:

$$L(M_i|data) \propto \exp\left\{-\frac{1}{2}\Delta_i(AIC)\right\}.$$

To calculate the AIC weights, $w_i(AIC)$, each of the relative likelihoods are normalised. This involves dividing each likelihood by the sum of the likelihoods of the set of models being compared.

$$w_i(AIC) = \frac{\exp\left\{-\frac{1}{2}\Delta_i(AIC)\right\}}{\sum_{k=1}^K \exp\left\{-\frac{1}{2}\Delta_k(AIC)\right\}}.$$

Once normalised the weights sum to one and can be interpreted as the probability that M_i is the best model. The ratio of two weights can be used to calculate how much more likely the model with the higher weight is the best model. The model with the higher weight is the numerator in this instance. Using AIC weights provides an indication of statistical confidence for the model with the lowest AIC value (Wagenmakers & Farrell, 2004).

Despite the theory of using AIC for model selection, care needs to be taken, as a model is only as good as its assumptions, irrespective of performance criteria. Due to the assumptions underlying the inclusion of the explanatory variables and associated interactions in the model, the model with the lowest AIC may not be the best model for explaining or predicting the response variable.

Bootstrapping: Pulling CPUE up by its Bootstraps

In the simplest nonparametric problems we do literally sample from the data, and a common initial reaction is that this is a fraud. In fact it is not. A wide range of statistical problems can be tackled this way (Davison & Hinkley, 1997).

Bootstrapping is a statistical resampling technique used to assess uncertainty (Davison & Kuonen, 2002). The technique relies on computing power so has become more popular as computing technology has advanced. The name given to this method and the theory which allows this method to be efficiently implemented using computers is due to the work of Bradley Efron in the late 1970s. The name ‘bootstrap’ originates from the idiom “to pull yourself up by your own bootstraps” which means to improve your situation by your own efforts without any outside assistance. The name echoes Tukey’s jackknife, developed two decades earlier and is intended to convey the self-help nature of the algorithm (Efron, 1979b). Tukey’s suggestion for naming Efron’s method was ‘shotgun’ due to Tukey’s view that the method could “blow the head off any problem if the statistician [could] stand the mess” (Efron, 1979a). Modern computing power has taken the mess out of the method and it has become a useful tool for estimating uncertainty.

Bootstrapping relies on samples of the original data to calculate estimators of interest and quantify the level of uncertainty. One of the key applications of bootstrapping is the estimation of confidence intervals to give an indication of the variance of point estimates such as the mean or median of a distribution (Orloff & Bloom, 2014). Implementation of bootstrapping for a dataset of size n involves the following steps:

1. Take a random sample of size n (the size of the original dataset) with replacement from the original data. This is the bootstrap sample.
2. Compute the statistic of interest on the bootstrap sample. This statistic is the bootstrap statistic.
3. Repeat Steps 1 and 2 k times to get k bootstrap samples. These k samples form a bootstrap distribution which can be used to calculate standard errors and confidence intervals to assess the variability of the statistic.

The magnitude of k is typically large and common choices include 1000, 10000 or 100000. The larger the value of k , the greater the accuracy of the variance estimates (Morgan, 2012). The choice of k is based on a combination of the required accuracy of the variance estimates and computational limitations.

Given that the random sample is taken with replacement, each of the bootstrap samples will be missing some data values and include repeats of others. In practise, the approximation of the bootstrap statistic may not be very accurate, but the approximation of the variance of this statistic is well-approximated by this method, especially for large values of k . The accuracy of the variance estimates from resampling the data k times is a result of the law of large numbers (Orloff & Bloom, 2014).

Bootstrapping is used here to analyse the behaviour of a selection of variables either side of the breakpoint. The means of the chosen variables either side of the breakpoint are compared using independent two sample t-tests. Bootstrapping, which employed a sample of 1000 was used in order to obtain confidence intervals for the pre and post breakpoint means.

When comparing CPUE either side of the breakpoint, each run of the bootstrap algorithm re-estimated the location of the breakpoint before calculating the means and confidence intervals. The breakpoint remained at 2004 for all of the bootstrap runs.

Cross Validation: To Explain or to Predict?

Linear models are a tool for explaining the relationship between a response variable and a selection of explanatory variables. This relationship can also be used for prediction. Model selection criteria must therefore take both the explanatory and predictive abilities of a model into account. AIC was derived based on predictive ability (Shmueli, 2010) but as outlined by the following quote, it is possible for this predictive ability to be confounded by the idiosyncrasies of the data.

Testing the procedure on the data that gave it birth is almost certain to overestimate performance, for the optimizing process that chose it from among many possible procedures will have made the greatest use possible of any and all of the idiosyncrasies of those particular data... As a result, the procedure will likely work better for these data than for almost any other data that will arrive in practice (Mosteller & Tukey, 1977).

Cross validation is used to assess the predictive ability of the LMs for CPUE as it provides a method of obtaining approximately unbiased estimators of prediction error (Efron & Gong, 1983). It is designed to identify the model with the best average predictive ability (Shao, 1993). There are two key cross validation techniques, known as leave-one-out or K -fold cross validation. The five key steps involved in both of these cross validation techniques are:

1. Remove a point or set of points from the dataset which form the test set.
2. Recalculate the prediction model with the remaining set of points which form the training set.
3. Predict the response variable and associated prediction error for the removed points based on the new model.
4. Repeat Steps 1- 3 until there is a prediction error for each point.
5. Average the prediction errors over all the points or sets of points removed.

Leave-one-out is asymptotically inconsistent in that the probability of selecting the model with the best prediction ability does not converge to one as the total number of observations approaches infinity, i.e. as $n \rightarrow \infty$ (Shao, 1993). On top of the potential inaccuracies with leave-one-out, it is also relatively computationally expensive (Kohavi, 1995) (Zhang, 1993). These inherent issues with the leave-one-out algorithm make K -fold cross validation, with a sensible choice of K , a preferable alternative.

K -fold cross validation involves randomly splitting the data into K mutually exclusive groups. The randomisation of groups is important to minimise possible bias (Zhang, 1993). If the full dataset can be viewed as a randomly drawn sample from the population of interest, then randomly selected folds for validation should be representative of a sample of future observations (Picard & Cook, 1984). The

model is fitted using the remaining $K-1$ groups and is then used to predict the response variable for the K th group. This process of excluding, fitting and predicting is repeated K times and the resulting prediction errors are averaged over the K folds.

The choice of K is a key component of K -fold cross validation. Small K are less computationally expensive and minimise the variance of the prediction errors. This also leads to an increase in bias. The choice of a larger K leads to a decrease in bias, but an increase in variance (Arlot, 2010). This leads to a trade-off between minimising both bias and variance.

Ten-fold cross validation is considered to be a better choice than leave-one-out (Kohavi, 1995). Simulations performed by Zhang found that $K \geq 5$ provides reliable estimates of prediction error. The most dramatic improvement occurs between $K = 2$ and $K = 10$. After $K = 10$, improvements are small and given the extra computational burden of increasing K , there is little advantage in going above 10 folds (Zhang, 1993). Thus, common choices of K are between 5 and 10 (Arlot, 2010). A large scale simulation experiment involving half a million runs with variations in number of folds and stratification within folds indicated that for real-world datasets, 10-fold stratified cross validation is the best method (Kohavi, 1995). The best choice is also dependent on the dataset, the purpose of the cross validation and the signal to noise ratio. A different choice of K may be appropriate in these circumstances (Arlot, 2010).

Based on suggestions from the literature, 10-fold cross validation was chosen for evaluating the LMs of CPUE. Stratification of the folds was considered such that the mean CPUE was approximately equal for each of the 10 folds, but it was decided that this extra step would merely increase the computation required and complicate the algorithm unnecessarily. The decision to use stratified cross validation is often more appropriate when the response variable is categorical.

There are many ways of calculating cross validation errors and the choice of the most suitable error is related to the model type, prediction method and whether the response variable is discrete, continuous or categorical. When conducting cross validation on CPUE, which is a continuous variable, the error is the difference between the predicted and the true value. The errors for each comparison are stored in a vector from which the overall prediction errors are calculated. Two appropriate ways of summarising this error are mean absolute error (MAE) and root mean squared error (RMSE). These two errors are complimentary in nature as using them both in conjunction provides a more informative picture of the error distribution (Chai & Draxler, 2014).

The formulas for MAE and RMSE are as follows:

$$MAE = \frac{1}{n} \sum_{i=1}^n |e_i|$$

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n e_i^2}.$$

Here, n is the number of data points involved in the cross validation and e_i is the error corresponding to the i th prediction which is calculated as follows:

$$e_i = (\text{predicted CPUE})_i - (\text{actual CPUE})_i.$$

The MAE is a measure of the accuracy of an estimator. It measures the average magnitude of the errors without considering their direction. It is a linear score, thus each error is given equal weighting in the final average. The RMSE incorporates the standard deviation of the estimator and its bias. If the error distribution is unbiased, the RMSE is equivalent to the standard error (SE). Squaring the errors before taking the average results in larger errors being given a higher weighting making the RMSE useful for identifying large errors.

It is common to use the MAE and RMSE in conjunction to get an indication of the variation in the errors. The RMSE will always be greater than or equal to the MAE. The larger the difference between the two errors, the greater the variance in the individual prediction errors. Ideally the MAE, RMSE and the difference between them will be small. The use of MAE and RMSE is summarised by Chai and Draxler (2014).

The results from the cross validation of the linear models for CPUE will give an indication of the models' ability to predict CPUE as well as to give estimates of the associated species abundance.

A Note on Boxplots

The boxplots included throughout the results chapter are Tukey boxplots where the band inside the box corresponds to the median and the bottom and top of the box correspond to the first and third quartiles respectively. The whiskers extend to 1.5 times the inter-quartile range and any data points outside this range are plotted as outliers with a small circle.

Results of Modelling CPUE

Exploratory Data Analysis

Exploratory analysis covered 17 years of orange roughy catches in the Priceless region with 99% of the fishing events occurring throughout a period of 14 years from September 1998 to July 2012 (Figure 15A). After initial exploratory analysis including all Priceless catches, it seemed pertinent to restrict the data from January 1998 to December 2012 to ensure that there were sufficient catches per year to allow for meaningful comparisons.

The average annual CPUE plots (Figure 15D - Figure 18D) are a coarse CPUE measure where effort for a given time period is defined as the frequency of fishing events in that period. The peak CPUE, based on frequency of fishing events, for the Priceless region occurred in 2002 (Figure 15D). This is followed by a sharp drop in CPUE for the years of 2003 and 2004, despite the number of fishing events peaking in 2004 (Figure 15B). After peaking in 2002, the annual CPUE for the Priceless region dropped and the trend is generally decreasing from 2005 onwards.

The North Pukaki region is spatially adjacent to Priceless so it is highly likely that the two stocks experience similar conditions and therefore have similar biological parameters. As a result, the two stocks are expected to have a similar response to fishing pressure.

North Pukaki was fished earlier than Priceless with light fishing pressure from 1989 onwards with the peak annual CPUE occurring in 1995, three years before any substantial effort in Priceless (Figure 16D). After the peak in 1995, the annual CPUE for North Pukaki exhibits a generally decreasing trend. The increased number of events and the decrease in annual CPUE are indicators of overfishing (Figure 16B, D).

The annual CPUE data for the Spawning Box and the NZ EEZ (Figure 17D and Figure 18D) tell a different story. There is an overall increasing trend for the Spawning Box orange roughy fishery (Figure 17D). Effort, defined as the number of fishing events, is highest in the period 1999-2005 which corresponds to a period of high yearly catch (Figure 17B, C). The number of events per year drops around 2004 with the number of events being approximately constant from 2006 onwards. The drop in catch is less than the decrease in effort resulting in the increasing annual CPUE (Figure 17D).

During the period from 2005 to 2014 the annual CPUE in the NZ EEZ was increasing (Figure 18D). By contrast, throughout this time period the annual CPUE for the Priceless region was decreasing (Figure 15D).

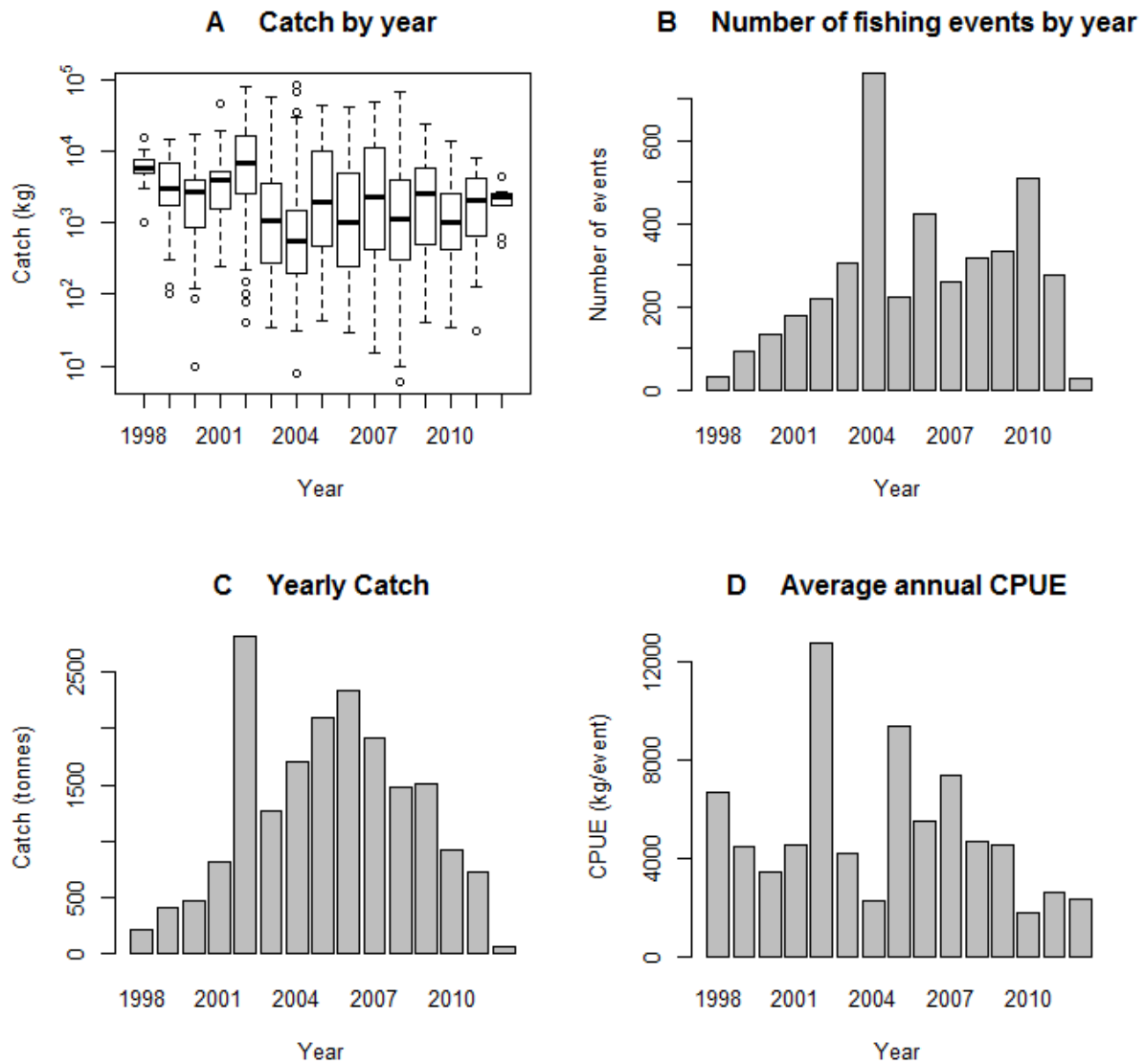


Figure 15 Exploratory data analysis for catch data from the Priceless region from 1998 - 2012. Plot A is a Tukey boxplot of catch (kg) in the Priceless region plotted by year. The number of fishing events per year are plotted in a bar graph in Plot B, with the peak number of events occurring in 2004. Tonnes are used as the units of yearly catch in Plot C which is a bar graph of the total catch weight caught in the Priceless region in the most active years of the fishery. The peak catch occurred in 2002, two years prior to the peak in number of events. The components of Plots B and C are combined to create Plot D which is the annual average CPUE where CPUE is measured as kg per event. The peak annual CPUE occurred in 2002 which is the year with the peak annual catch (Plot C). 2004, the year with the peak number of fishing events (Plot B) has one of the lowest annual CPUE values of the series. There is a general decreasing trend in annual CPUE from 2005 onwards (Plot D) which is similar to the decreasing trend seen in annual catch from 2006 onwards (Plot C).

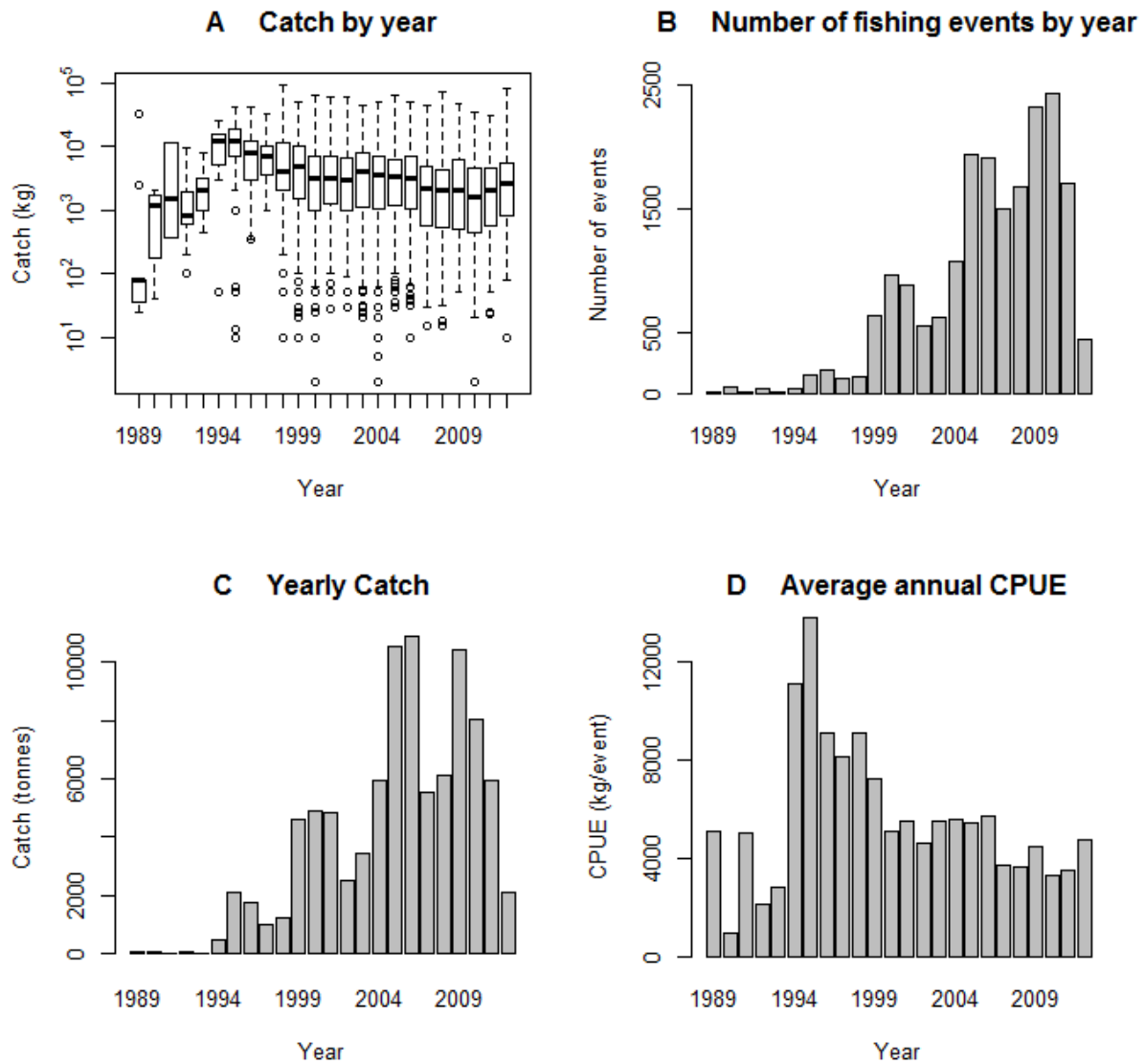


Figure 16 Exploratory data analysis for catch data from the North Pukaki region from 1989 - 2012. The time series of catch data for North Pukaki begins in 1989 with very few fishing events occurring in the exploratory phase prior to 1994. The number of fishing events has a generally increasing trend over time (Plot B) which is approximately mirrored by catch (in tonnes) (Plot C). The annual CPUE is not as promising (Plot D). The peak annual CPUE occurred in 1995 and past this point there is a generally decreasing trend in annual CPUE.

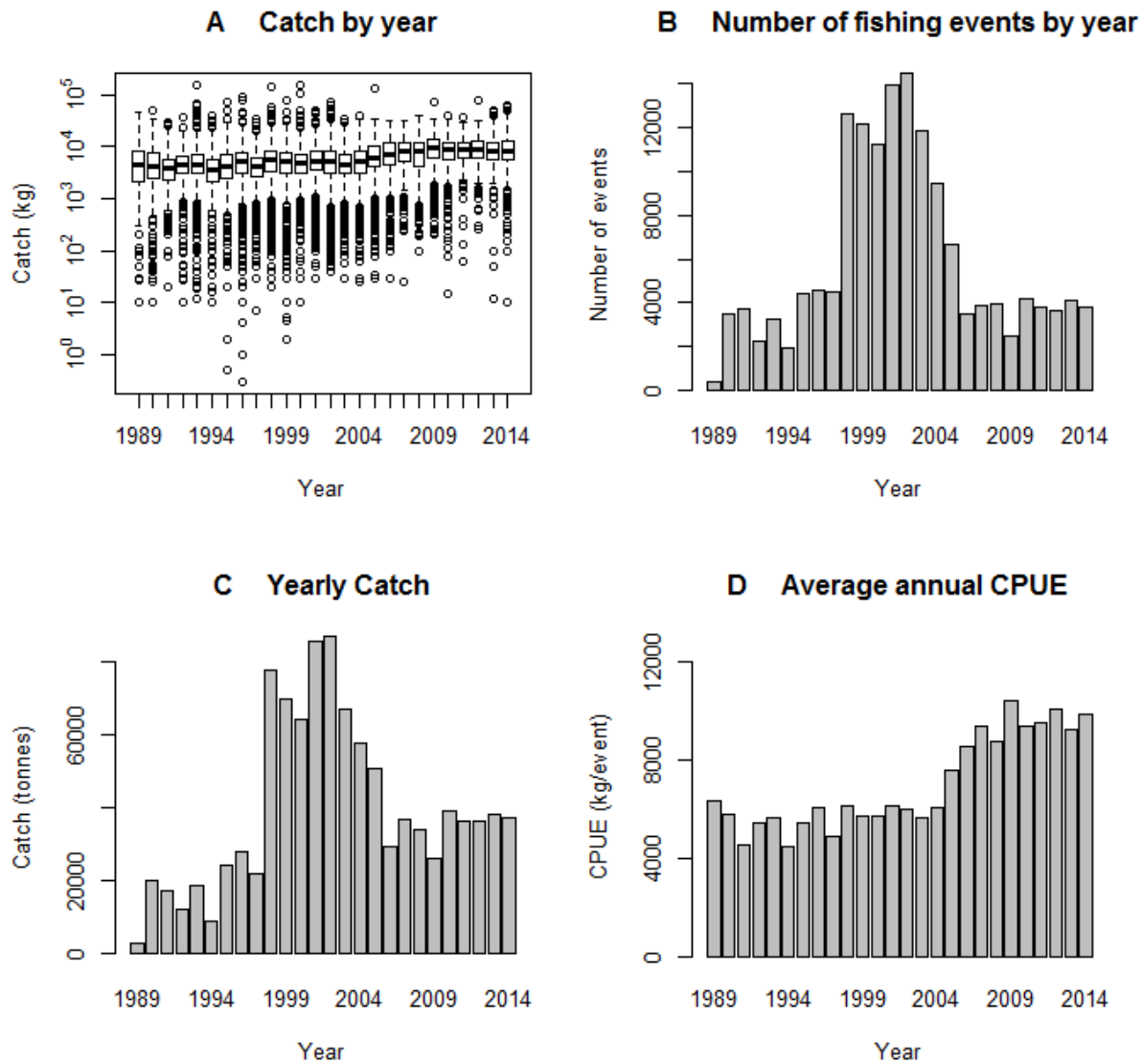


Figure 17 Exploratory data analysis for catch data from the Spawning Box region in the East and South Chatham Rise from 1989 - 2014. The boxplot of catch by year demonstrates the high level of variability between individual catches (Plot A). The number of annual fishing events (Plot B) and the yearly catch (Plot C) exhibit very similar behaviour and share a peak in 2002. The average annual CPUE for the Spawning Box reaches its peak seven years later in 2009 (Plot D). There is a generally increasing trend in the annual CPUE which contrasts to the trends seen in annual CPUE in the Priceless and Pukaki regions (Figure 15D and Figure 16D).

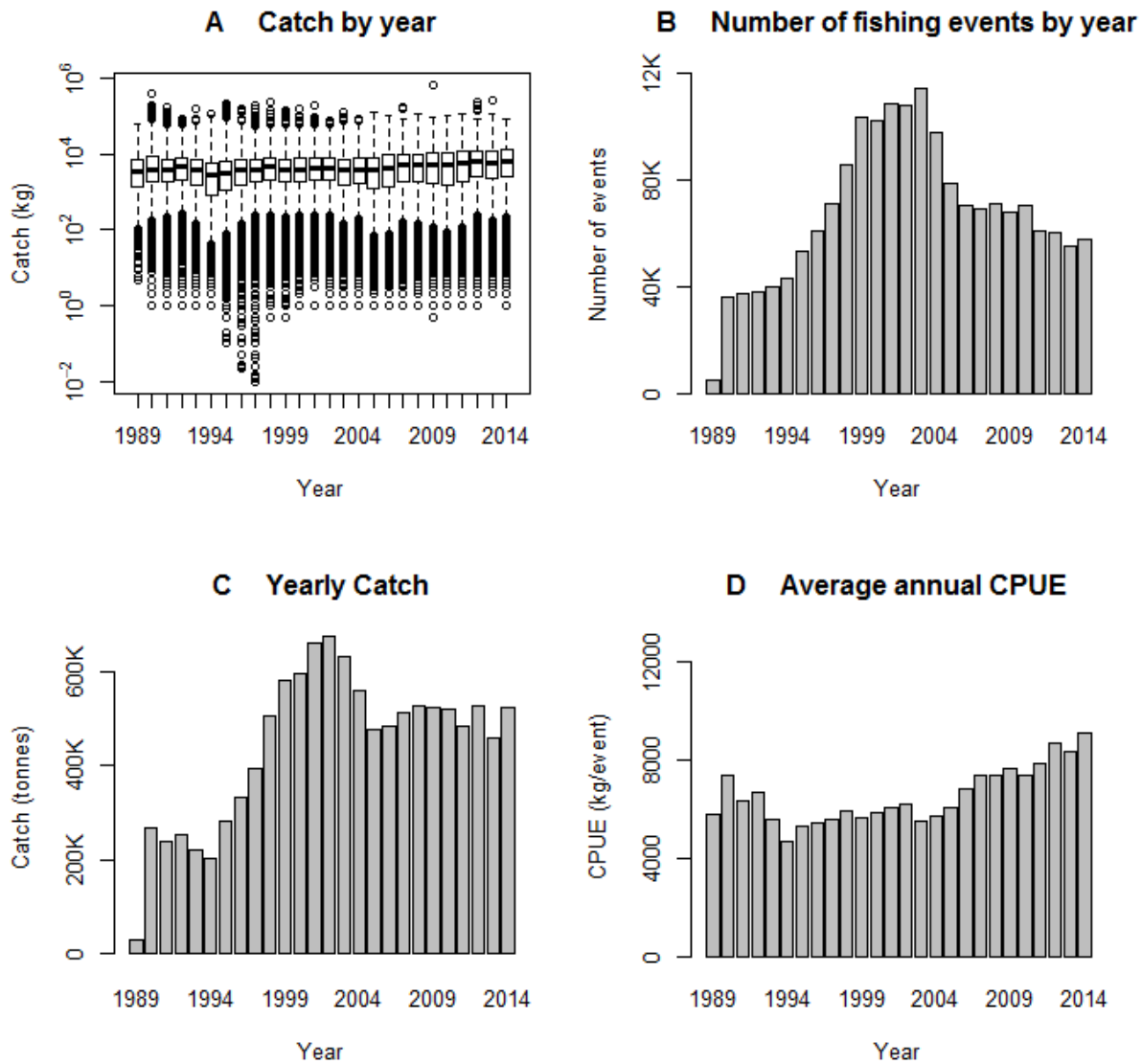


Figure 18 Exploratory data analysis for catch data from the NZ EEZ from 1989 - 2014. The variation in catch in the NZ EEZ is greater than the variation seen in the three other regions which is expected due to the larger volume of catches included (Plot A). The peak number of fishing events occurs in 2003 (Plot B), one year after the peak in yearly catch (Plot C). The peak in annual CPUE occurred in 2014, the final year in the series (Plot D). Annual CPUE was high in the initial years and from 1994 onwards exhibits a generally increasing trend.

When comparing Figure 15 - Figure 18, note that Plots B and C have different scales on the y axes as the different sizes of each stock leads to the number of fishing events and annual catch weights having vastly different magnitudes. The scaling of catch to create the annual CPUE plots (Plot D) leads to each region having a similar range for CPUE, allowing each of the four regions to have the same y axes for this plot.

Standardising CPUE using Linear Models

Assumptions of the catch data from the Priceless region and expertise from published research on CPUE led to the fitting and comparison of a selection of models for CPUE. CPUE is dynamic and this property gives rise to the importance of modelling with temporal variables. The most applicable temporal scales for modelling CPUE were identified in the preliminary phase of modelling. The temporal scale was explored at daily, monthly and yearly levels. To explore the annual seasonality of CPUE, a selection of models with varying numbers of cycles were fitted to mimic seasonal effects (Table 8).

A selection of LMs were fitted in a forward stepwise manner to find the best model for explaining CPUE (Table 9 - Table 12). The tables to follow include descriptions to aid practical interpretation of each model and comparisons were made using model AIC and the corresponding weights. The lowest AIC in each table is highlighted in bold.

The best model in Table 9 was an additive model consisting of categorical month and year. This model allows abrupt changes in CPUE for each unit of time. Table 9 is used to explore additive models with a selection of temporal variables and does not include any interactions between the year and month terms implying that the monthly pattern in CPUE is equivalent from year to year. The models in Table 10 expand on those in Table 9 with the inclusion of interactions. Including these interactions assumes that there is inter-year variance of the monthly pattern in CPUE.

+			Year		
			NULL Annual CPUE remains constant	Year_F CPUE fluctuates annually	Year_N An annual linear trend in CPUE
The species variable is not included in these models indicating that CPUE behaves the same for each species	Month	NULL Monthly CPUE remains constant	19652 (904, 0)	18874 (126, 0)	19526 (778, 0)
		Month_F Monthly fluctuations that repeat annually	19289 (541, 0)	18748 (0, 0.87)	19073 (325, 0)
		Month_N A linear trend within each year which repeats annually	19588 (840, 0)	18847 (99, 0)	19388 (640, 0)
		Month_N1 A seasonal pattern with one peak and trough	19512 (764, 0)	18806 (58, 0)	19274 (526, 0)
		Month_N2 A seasonal pattern with two peaks and troughs	19515 (767, 0)	18792 (44, 0)	19240 (492, 0)
		Month_N3 A seasonal pattern with three peaks and troughs	19398 (650, 0)	18767 (19,0)	19155 (407, 0)
		Month_N4 A seasonal pattern with four peaks and troughs	19337 (589, 0)	18757(9,0.0097)	19110 (362, 0)
		Month_N5 A seasonal pattern with five peaks and troughs	19306 (558, 0)	18752 (4, 0.12)	19079 (331, 0)

Table 9 AIC values and model interpretations for CPUE models with variants on year and month as explanatory variables. The ΔAIC and AIC weights are included in brackets to the right of the model AIC. The model with the lowest AIC is highlighted in bold and implies abrupt changes in monthly and annual CPUE.

*			Year	
			Year_F CPUE fluctuates annually	Year_N An annual linear trend
The species variable is not included in these models indicating that CPUE behaves the same for each species	Month	Month_F Monthly CPUE fluctuates	18688 (16, 0.00033)	19011 (339, 0)
		Month_N Linear trends which vary annually	18838 (166, 0)	19381 (709, 0)
		Month_N1 A seasonal pattern with one peak and trough	18750 (78, 0)	19250 (578, 0)
		Month_N2 A seasonal pattern with two peaks and troughs	18684 (12, 0.0025)	19210 (538, 0)
		Month_N3 A seasonal pattern with three peaks and troughs	18672 (0, 0.98)	19052 (380, 0)
		Month_N4 A seasonal pattern with four peaks and troughs	18680 (8, 0.018)	19031 (359, 0)
		Month_N5 A seasonal pattern with five peaks and troughs	18688 (16, 0.00033)	19017 (345, 0)

Table 10 AIC values for CPUE models with variants on year and month and their interactions as explanatory variables. The Δ AIC and AIC weights are included in brackets to the right of the model AIC. The model with the lowest AIC is in bold and models CPUE using month with three seasonal harmonics which vary annually and year with abrupt changes.

The AIC values are significantly lower for the models that include interactions which confirms the high temporal variability in CPUE. Including interactions between year and month leads to a best interactive model with different explanatory variables to the best additive model (Table 9 and Table 10). The best model of those with interactions still implies CPUE is subject to abrupt fluctuations between years while the inclusion of a monthly harmonic term with three peaks and troughs corresponds to seasonal fluctuations in CPUE (Table 10). The interaction indicates that the pattern of these seasonal fluctuations changes over time. The Δ AIC for the best temporal models is 76 (Table 9 and Table 10). The magnitude of this difference indicates that the interactions between year and month are significant in explaining the variation within CPUE.

<div>+</div> <div>three_species</div>			Year		
			<div>NULL</div> <div>Annual CPUE remains constant</div>	<div>Year_F</div> <div>CPUE fluctuates annually</div>	<div>Year_N</div> <div>An annual linear trend in CPUE</div>
Adding species to the model implies that each species has the same CPUE pattern but with different magnitudes	Month	<div>NULL</div> <div>Monthly CPUE remains constant</div>	17486 (568, 0)	17013 (95, 0)	17478 (560, 0)
		<div>Month_F</div> <div>Monthly fluctuations which repeat annually</div>	17221 (303, 0)	16918 (0, 0.99)	17165 (247, 0)
		<div>Month_N</div> <div>A linear trend within each year which repeats annually</div>	17425 (507, 0)	16996 (78, 0)	17395 (477, 0)
		<div>Month_N1</div> <div>A seasonal pattern with one peak and trough</div>	17374 (456, 0)	16970 (52, 0)	17332 (414, 0)
		<div>Month_N2</div> <div>A seasonal pattern with two peaks and troughs</div>	17359 (441, 0)	16955 (37, 0)	17286 (368, 0)
		<div>Month_N3</div> <div>A seasonal pattern with three peaks and troughs</div>	17284 (366, 0)	16937 (19, 0)	17223 (305, 0)
		<div>Month_N4</div> <div>A seasonal pattern with</div>	17257 (339, 0)	16932(14,0.0009)	17197 (279, 0)
		<div>Month_N5</div> <div>A seasonal pattern with five peaks and troughs</div>	17234 (316, 0)	16927 (9, 0.011)	17173 (255, 0)

Table 11 AIC values for CPUE models with species and variants on year and month as explanatory variables. The Δ AIC and AIC weights are included in brackets to the right of the model AIC. The model with the lowest AIC is in bold and includes species and year and month as categorical variables. This model implies that a change in species, year or month corresponds to an abrupt change in CPUE.

* three_species			Year	
			Year_F CPUE fluctuates annually	Year_N An annual linear trend
The interaction with species implies that each species has its own CPUE pattern	Month	Month_F Monthly CPUE fluctuates	16102 (9, 0.011)	16901 (808, 0)
		Month_N Linear trends which vary annually	16462 (369, 0)	17335 (1242, 0)
		Month_N1 A seasonal pattern with one peak and trough	16297 (204, 0)	17236 (1143, 0)
		Month_N2 A seasonal pattern with two peaks and troughs	16117 (24, 0)	17129 (1036, 0)
		Month_N3 A seasonal pattern with three peaks and troughs	16093 (0, 0.98)	16999 (906, 0)
		Month_N4 A seasonal pattern with four peaks and troughs	16121 (28, 0)	16926 (833, 0)
		Month_N5 A seasonal pattern with five peaks and troughs	16102 (9, 0.011)	16914 (821, 0)

Table 12 AIC values for CPUE models with species and variants on year and month as explanatory variables including interactions. The ΔAIC and AIC weights are included in brackets to the right of the model AIC. The model with the lowest AIC is in bold and includes species, month with three harmonics, categorical year and the full set of interactions.

Given that species is an important descriptor of CPUE, the species variable is incorporated into the LMs (Table 11 and Table 12). This is biologically sensible as different species are likely to display different patterns of aggregation and to occupy different spatial features, which is likely to have an effect on catchability.

The best additive model including species implies that CPUE is different for each of the three species groupings and that CPUE experiences abrupt fluctuations between years and months (Table 11). The ΔAIC of 1830 for the best additive models demonstrates that the species caught informs CPUE (Table 9 and Table 11).

Given the likelihood of annual and seasonal changes in the catchability of each species in relation to abundance, recruitment and environmental factors, an interaction between species and the temporal variables is expected to yield beneficial results (Table 12).

The model with the lowest AIC value includes a unique CPUE for each species group, month with three harmonics, corresponding to a seasonal pattern of three peaks and troughs, and year as a categorical variable representing abrupt changes in CPUE between years (Table 12). There are two other models in Table 12 with non-zero AIC weights. The first of these is the model where each species has a unique CPUE and changes in month or year result in abrupt fluctuations in CPUE. The second of these includes a unique CPUE for each of the three species groups, monthly seasonality consisting of five peaks and troughs and year as a categorical variable representing abrupt changes in CPUE between years. The latter of these models has the same descriptive accuracy as the former, but is more parsimonious. The more parsimonious model is a better candidate for model prediction.

SUMMARY:

Table 9 - Table 12 give an extensive overview of models for CPUE involving species and a selection of temporal variables. Based on assumptions of fishing dynamics, the most relevant model allows for each species to have its own time series of CPUE and this time series involves yearly and monthly patterns which experience annual variability.

Standardising CPUE using Environmental and Effort Variables

There are many other variables which have been known to affect CPUE. These include environmental conditions and effort variables relating to fishing equipment. Linear models were fitted with a selection of environmental and effort variables to determine their influence on CPUE.

The effort and environmental variables are added to the model including species, seasonal month with three peaks and troughs, year with abrupt changes in CPUE, and the full set of interactions which is the most descriptive of the temporal models (Table 12). Of the effort and environmental variables modelled, one of them has a smaller AIC value than that of any of the previous models that include species, temporal variables and the associated interactions (Table 13). Modelling CPUE with sea surface temperature leads to the lowest model AIC of any of the models fitted so far and a ΔAIC of 2439 between this model and the next best model (Table 12 and Table 13). The estimate of the coefficient for surface temperature is significant and suggests that a 1°C increase in temperature is likely to result in an increase in CPUE of 1.3kg/h.

The model containing species, surface temperature and the interaction between them has a higher AIC than the model which also includes the temporal variables, month with three harmonics and year with abrupt fluctuations. The most descriptive model up to this point includes all four of the aforementioned variables and their interactions (Table 13).

Explanatory variables	Model AIC	ΔAIC	AIC weights
three_species * month_N3* year_F * surface_temp	10862	0	1
three_species * month_N3* year_F + surface_temp	11018	156	0
three_species * surface_temp	11900	1038	0
three_species * month_N3* year_F + bottom_temp	15448	4586	0
three_species * month_N3* year_F + effort_speed	16007	5145	0
three_species * month_N3* year_F + effort_width	16054	5192	0
three_species * month_N3* year_F + bottom_depth	16054	5192	0
three_species * month_N3* year_F + effort_height	16072	5210	0
three_species * month_N3* year_F + effort_depth	16074	5212	0

Table 13 AIC values for CPUE models with a selection of effort and environmental variables. The adjacent columns include the ΔAIC and AIC weights to facilitate comparison of model performance. The best model is in bold and includes species, month with three harmonics, year as a categorical variable, surface temperature and the corresponding interactions between these variables.

After discovering the descriptive ability of surface temperature, the next step was to create a series of plots to gain more insight about the variable. The initial concern was the possibility that surface temperature was linked to locations or accuracy of measurement from individual vessels which would confound any results involving this variable.

The relationship between surface temperature and location is explored (Figure 19). Surface temperatures are grouped into one of three buckets which cover the full range of the variable. Each of these buckets is then assigned a different colour where yellow, orange and red correspond respectively to low, medium or high sea surface temperatures for the Priceless region. Each bucket represents a temperature range of 5°C. These temperatures are then plotted at their measurement location using the start latitude and longitude coordinates for a given trawl.

The spread of points in all three of the temperature buckets leads to the conclusion that there is no obvious spatial pattern in sea surface temperature which would confound the results of the CPUE models (Figure 19). The variation in sea surface temperature is best attributed to natural variation from seasonal and climatic components.

Given the possibility of between vessel differences in equipment for measuring temperature, it was considered necessary to investigate whether there were any obvious inadequacies in surface temperature measurements for the 14 vessels. A combination of graphing catch locations by vessel and inspection of the mean and range of surface temperature for each vessel did not pick up any obvious inconsistencies which would question the suitability of modelling CPUE with surface temperature.

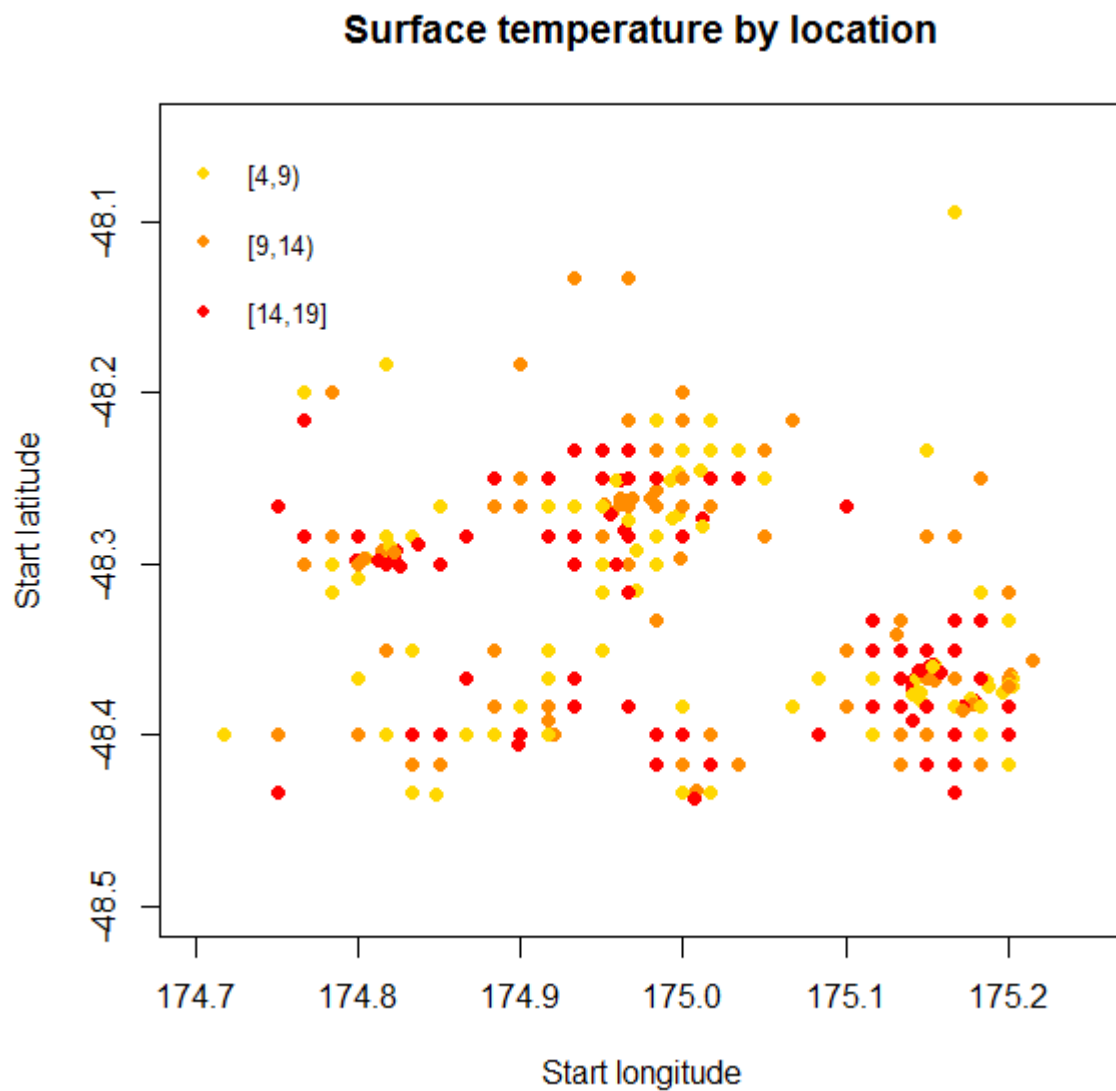


Figure 19 An exploration of sea surface temperature by location. The sea surface temperature recorded in the dataset ranges from 4.5 – 19 °C. This temperature range has been divided into three buckets, each of which span 5°C. The temperature gradient is plotted using the coordinates of the starting location of each trawl in the Priceless region. The three colours, yellow, orange and red correspond to a sea surface temperature measure in one of the three buckets, where the redder the colour, the warmer the temperature.

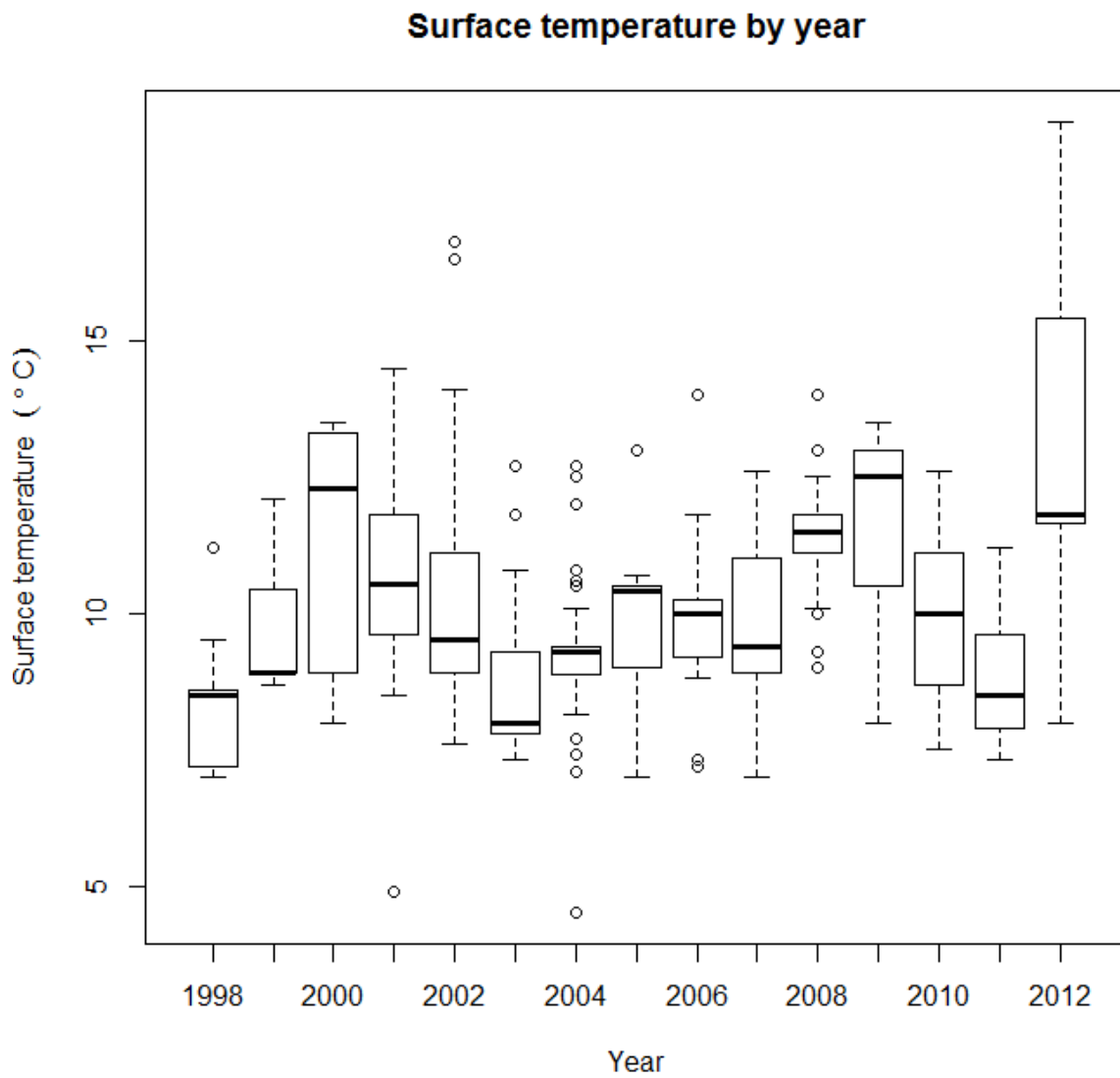


Figure 20 Time series of annual sea surface temperature measured in °C. A cyclic pattern can be seen in sea surface temperature prompting investigation into whether modelling CPUE with the SOI as an explanatory variable could lead to further insights and greater accuracy.

The cyclic nature of the surface temperature annual time series suggests the variable is driven by an underlying process (Figure 20). El Niño and La Niña phases are well-known to affect ocean temperatures and occur on average every two to seven years. The identification of a cycle within annual sea surface temperature led to modelling using the Southern Oscillation Index (SOI). A set of LMs were fitted to explore the ability of the SOI and a categorical variable corresponding to El Niño conditions to explain variation in CPUE (Table 14).

Explanatory variables	Model AIC	Δ AIC	AIC weights
three_species * month_N3* year_F + surface_temp	11018	0	1
three_species * month_N3* year_F + EN_LN	16092	5074	0
three_species * month_N3* year_F	16093	5075	0
three_species * month_N3* year_F + SOI	16095	5077	0

Table 14 AIC values for models of CPUE using the SOI and a categorical variable corresponding to El Niño conditions as explanatory variables. These two variables have been added to the best temporal model consisting of species, month with three harmonics and abrupt changes in in yearly CPUE (Table 12). This model and the model including surface temperature have been included in the table for comparison. The Δ AIC and AIC weights indicate that surface temperature cannot be adequately replaced by either the SOI or El Niño variables.

Adding the SOI or El Niño variables to the model for CPUE does not provide a significant improvement on the model which excludes them. If the model with surface temperature is not taken into account when calculating the AIC weights, there is a 55% chance that the model including the El Niño conditions is the best model and a 33% chance that the model without either SOI or El Niño is the best (Table 14). In fact, adding SOI to the model leads to a decrease in model performance.

SUMMARY:

Neither the SOI nor El Niño variables come remotely close to having the explanatory ability of surface temperature, despite their similarities and SOI being a significant explanatory variable for surface temperature. Thus, in order to take environmental conditions into account when modelling CPUE, the best variable available to do this is surface temperature.

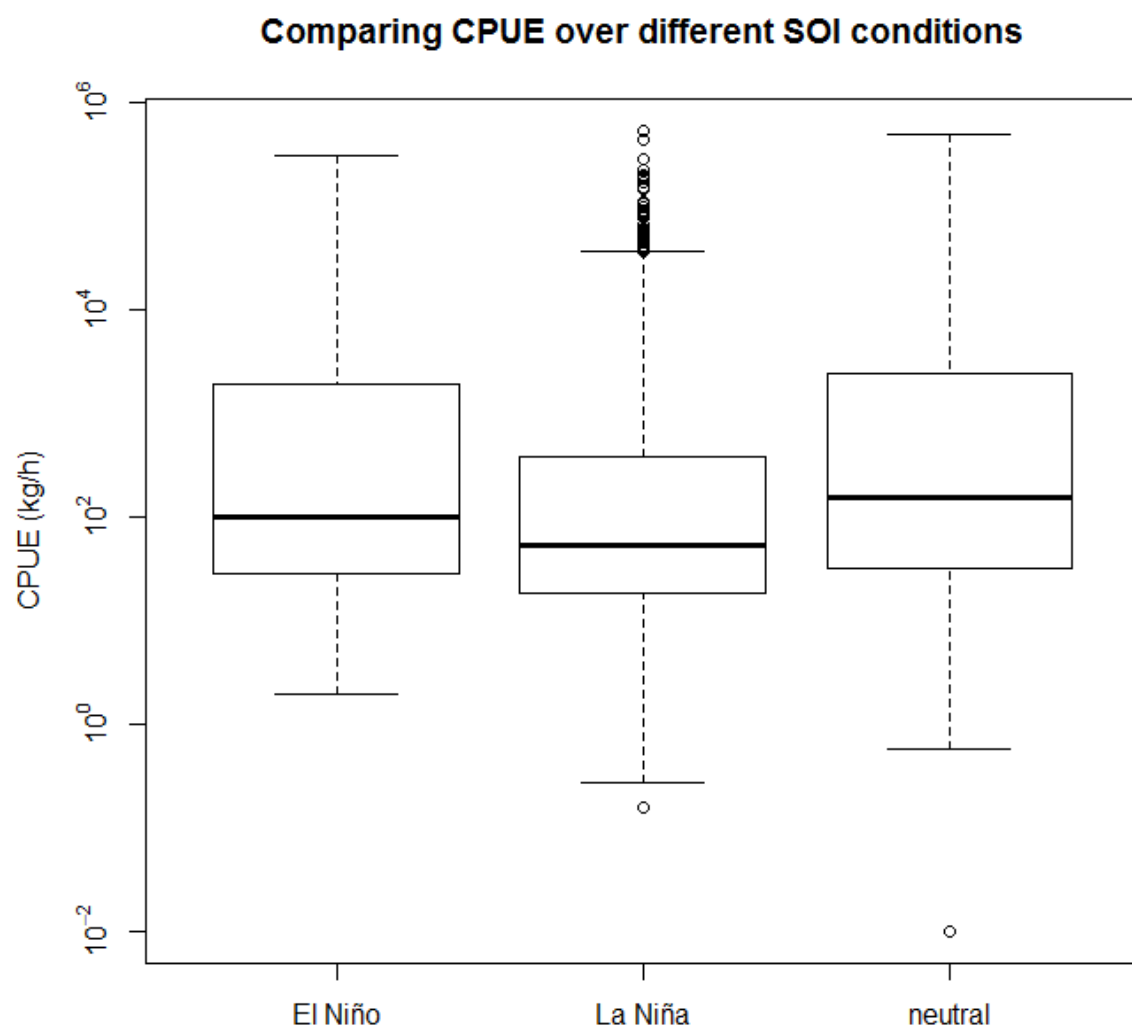


Figure 21 A comparison of CPUE for El Niño, La Niña and neutral conditions. The El Niño and neutral conditions have a similar effect on CPUE, whereas the La Niña conditions result in a lower CPUE overall.

Breakpoint Analysis

Inspection of the plot of the time series of CPUE by year leads to an interesting observation. The behaviour of the series in the early years of the fishery exhibits different behaviour to that in the latter years, implying the presence of a breakpoint around the year 2003-4 (Figure 22).

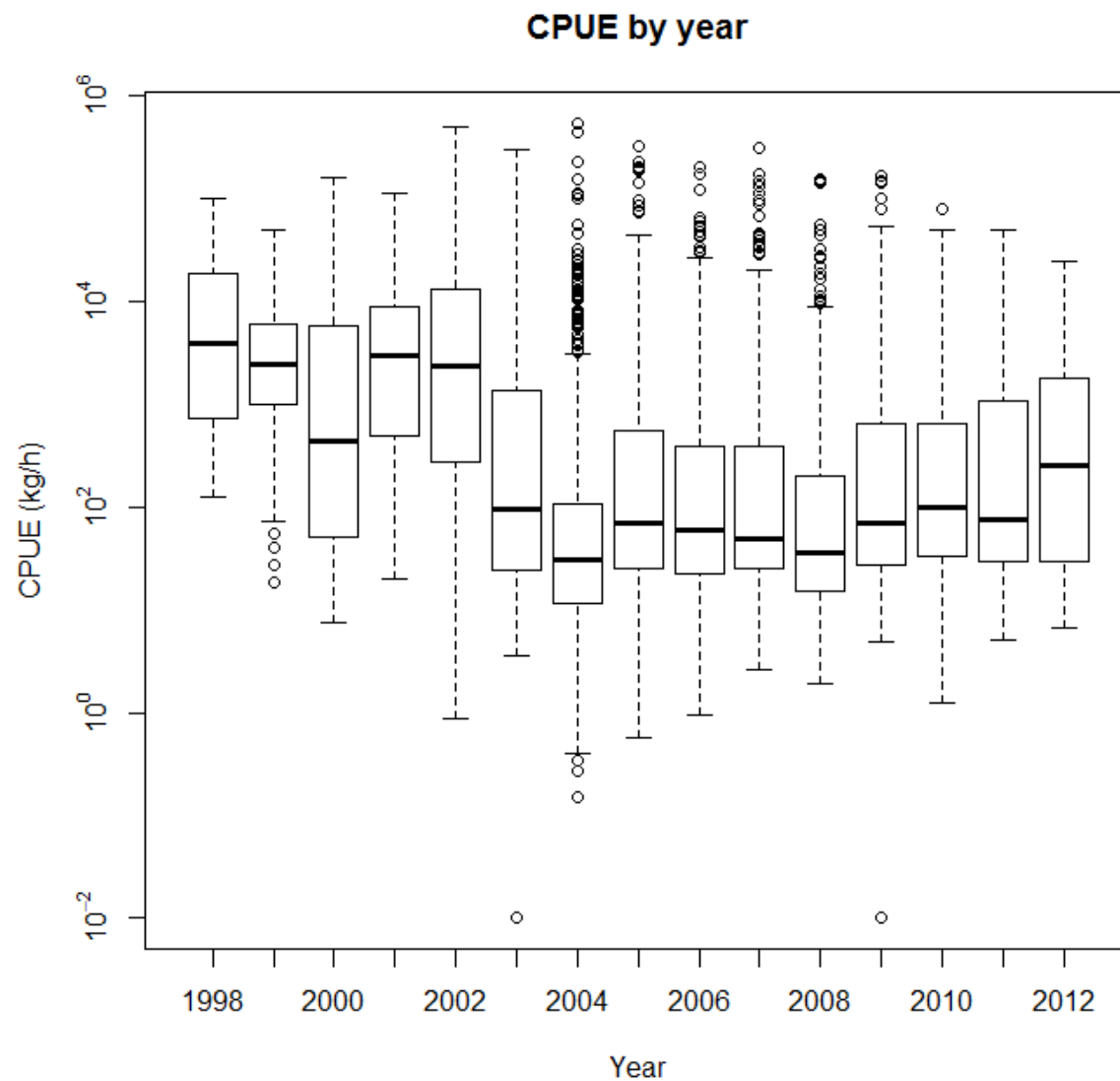


Figure 22 CPUE plotted by year with an apparent breakpoint around 2004. This suggests that an LM that includes an annual linear trend that incorporates a breakpoint could lead to an improved explanation of CPUE.

Visual inspection of Figure 22 suggests the location of a breakpoint in the CPUE measure. In order to quantify this breakpoint, fifteen potential models were fitted, allowing for each of the fifteen years in the dataset to be tested as candidate breakpoints. The AIC values of these fifteen candidate models were then compared to determine the best year to use for the breakpoint (Figure 23).

The year with the lowest AIC is the most suitable breakpoint for the CPUE data. In this case, the best choice of breakpoint is 2004. As the year decreases or increases either side of 2004, the model AIC values increase which provides support for the identification of 2004 as the optimal breakpoint. The Δ AICs for the candidate breakpoint models are all greater than 80. This gives the model with 2004 as the breakpoint an AIC weight of one and the remaining candidate models an AIC weight of zero.

After identifying the optimal breakpoint, a selection of LMs were fitted with a linear trend in year with a breakpoint at 2004. This new breakpoint model is fitted with all of the month terms including the harmonics (Table 15). The extension of these models involves the inclusion of interactions and the species variable (Table 15 - Table 18). A direct comparison can be made between these models and those in Table 9 to Table 12 which treat CPUE as undergoing abrupt annual fluctuations or having a linear trend with a constant slope.

SUMMARY:

Modelling CPUE with a yearly linear trend that incorporates a breakpoint leads to significantly better models than modelling with a constant linear trend. However, modelling CPUE with abrupt annual fluctuations leads to a model which is significantly better than those which incorporate a breakpoint in the linear trend. This indicates that although the presence of a breakpoint is significant, the annual variability in CPUE is better explained by a model allowing abrupt annual fluctuations.

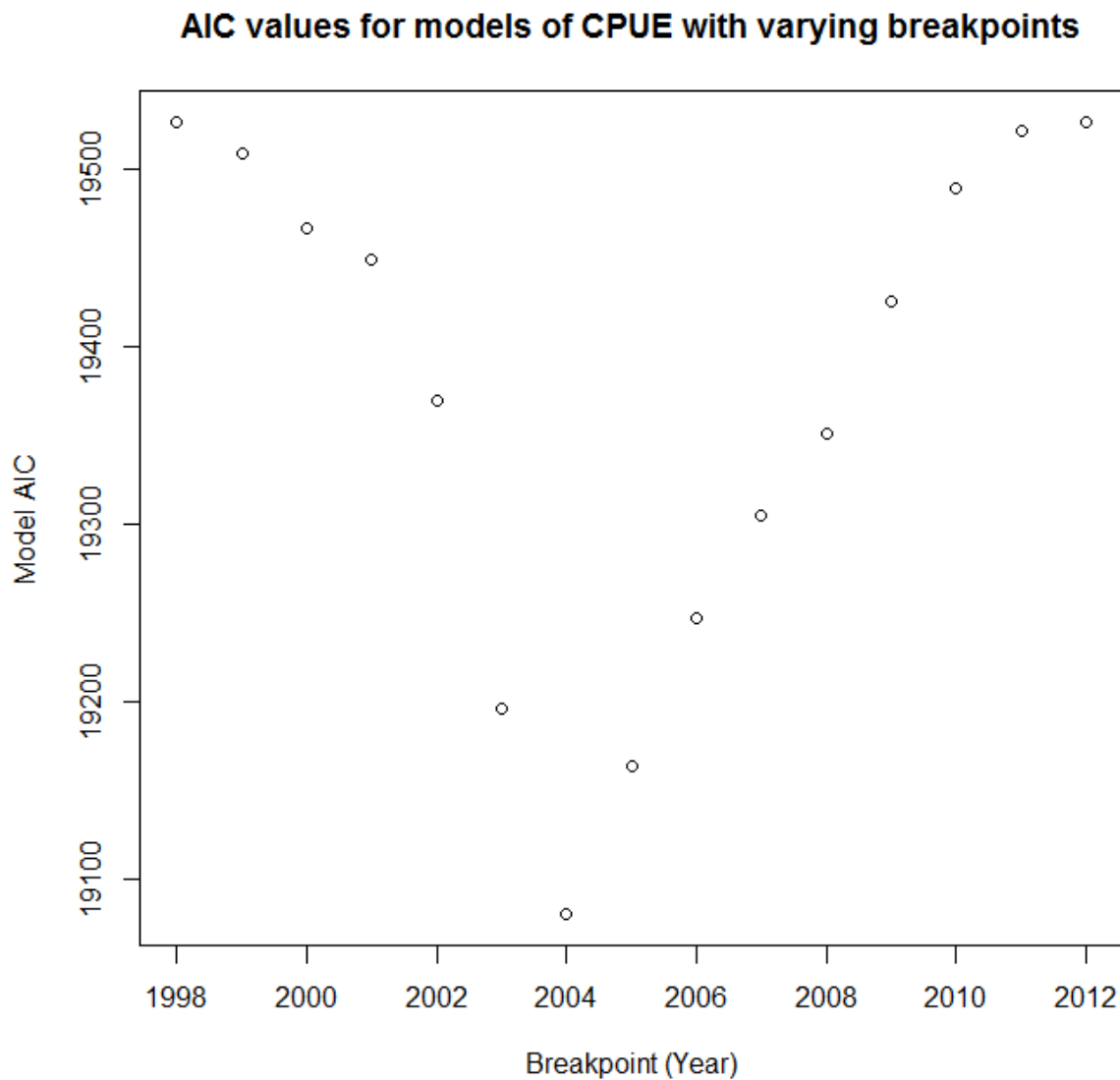


Figure 23 A plot of AIC values for candidate breakpoint years for modelling CPUE with a segmented linear trend. The breakpoint models are LMs fitted with an annual linear trend in CPUE where the slope of the trend is allowed to differ either side of the breakpoint. Fifteen separate models are fitted with the candidate breakpoint ranging from 1998 to 2012. The minimum AIC leads to identification of the year 2004 as the optimal breakpoint.

+		Annual linear trend with a breakpoint in 2004 $\alpha + \beta_1 year + \beta_2 (year - 2004) I_{\{year \geq 2004\}}$ CPUE undergoes an annual linear decrease where the rate of decrease is different either side of 2004.	
Species is not included in these models indicating that the CPUE behaves the same for each species	Month	NULL A monthly pattern for CPUE is absent	19080 (148, 0)
		Month_F A monthly pattern in CPUE results in fluctuations which repeat annually	18932 (0, 0.88)
		Month_N There is a linear trend in CPUE within each year which repeats annually	19070 (138, 0)
		Month_N1 Within each year there is a seasonal pattern for CPUE with one peak and trough which repeats annually	19015 (83, 0)
		Month_N2 Within each year there is a seasonal pattern for CPUE with two peaks and troughs which repeats annually	18998 (66, 0)
		Month_N3 Within each year there is a seasonal pattern for CPUE with three peaks and troughs which repeats annually	18948 (16, 0.0003)
		Month_N4 Within each year there is a seasonal pattern for CPUE with four peaks and troughs which repeats annually	18938 (6, 0.044)
		Month_N5 Within each year there is a seasonal pattern for CPUE with five peaks and troughs which repeats annually	18937 (5, 0.072)

Table 15 AIC values for models of CPUE which incorporate a linear trend with a breakpoint in 2004 and variants of month as explanatory variables. Variable interactions and species are not included in these models. The ΔAIC and AIC weights are included in brackets to the right of the model AIC. The best of these breakpoint models is in bold and includes abrupt monthly fluctuations in CPUE.

*		Annual linear trend with a breakpoint in 2004 $\alpha + \beta_1 year + \beta_2 (year - 2004) I_{\{year \geq 2004\}}$ CPUE undergoes an annual linear decrease where the rate of decrease is different either side of 2004.	
Species is not included in these models indicating that the CPUE behaves the same for each species	Month	Month_F A monthly pattern in CPUE results in fluctuations which vary annually	18792 (0, 0.69)
		Month_N There is a linear trend in CPUE within each year which varies annually	19036 (244, 0)
		Month_N1 Within each year there is a seasonal pattern for CPUE with one peak and trough which varies annually	18922 (130, 0)
		Month_N2 Within each year there is a seasonal pattern for CPUE with two peaks and troughs which varies annually	18875 (83, 0)
		Month_N3 Within each year there is a seasonal pattern for CPUE with three peaks and troughs which varies annually	18809 (17, 0.00014)
		Month_N4 Within each year there is a seasonal pattern for CPUE with four peaks and troughs which varies annually	18794 (2, 0.25)
		Month_N5 Within each year there is a seasonal pattern for CPUE with five peaks and troughs which varies annually	18797 (5, 0.057)

Table 16 AIC values for models of CPUE which incorporate a linear trend with a breakpoint in 2004 and variants of month as explanatory variables including interactions. The ΔAIC and AIC weights are included in brackets to the right of the model AIC. The model with the lowest AIC is in bold and includes abrupt monthly changes in CPUE. The model including month with four harmonics has the second lowest AIC with a weight of 0.25 making it a candidate for the best model in this table.

+ three_species		Annual linear trend with a breakpoint in 2004 $\alpha + \beta_1 year + \beta_2 (year - 2004) I_{\{year \geq 2004\}}$ CPUE undergoes an annual linear decrease where the rate of decrease is different either side of 2004.	
Adding species to the model implies that each species has the same CPUE pattern but with different magnitudes	Month	NULL A monthly pattern for CPUE is absent	17191 (120, 0)
		Month_F A monthly pattern in CPUE results in fluctuations which repeat annually	17071 (0, 0.96)
		Month_N There is a linear trend in CPUE within each year which repeats annually	17182 (111, 0)
		Month_N1 Within each year there is a seasonal pattern for CPUE with one peak and trough which repeats annually	17142 (71, 0)
		Month_N2 Within each year there is a seasonal pattern for CPUE with two peaks and troughs which repeats annually	17124 (53, 0)
		Month_N3 Within each year there is a seasonal pattern for CPUE with three peaks and troughs which repeats annually	17085 (14, 0.00087)
		Month_N4 Within each year there is a seasonal pattern for CPUE with four peaks and troughs which repeats annually	17080 (9, 0.011)
		Month_N5 Within each year there is a seasonal pattern for CPUE with five peaks and troughs which repeats annually	17078 (7, 0.029)

Table 17 AIC values for models of CPUE which incorporate a linear trend with a breakpoint in 2004 and variants of month and species as explanatory variables. The ΔAIC and AIC weights are included in brackets to the right of the model AIC. The model with the lowest AIC is in bold and includes abrupt monthly changes in CPUE. The models including month with three, four or five harmonics also have non-zero weights.

*		Annual linear trend with a breakpoint in 2004	
three_species		$\alpha + \beta_1 year + \beta_2 (year - 2004) I_{\{year \geq 2004\}}$ <p>CPUE undergoes an annual linear decrease where the rate of decrease is different either side of 2004</p>	
The interaction with species implies that each species has its own CPUE pattern	Month	Month_F A monthly pattern in CPUE results in fluctuations which vary annually	16553 (0, 0.81)
		Month_N There is a linear trend in CPUE within each year which varies annually	17013 (460, 0)
		Month_N1 Within each year there is a seasonal pattern for CPUE with one peak and trough which varies annually	16817 (264, 0)
		Month_N2 Within each year there is a seasonal pattern for CPUE with two peaks and troughs which varies annually	16693 (140, 0)
		Month_N3 Within each year there is a seasonal pattern for CPUE with three peaks and troughs which varies annually	16582 (14, 0)
		Month_N4 Within each year there is a seasonal pattern for CPUE with four peaks and troughs which varies annually	16556 (3, 0.18)
		Month_N5 Within each year there is a seasonal pattern for CPUE with five peaks and troughs which varies annually	16564 (11, 0.0033)

Table 18 AIC values for models of CPUE which incorporate a linear trend with a breakpoint in 2004 and variants of month and species as explanatory variables including interactions. The Δ AIC and AIC weights are included in brackets to the right of the model AIC. The model with the lowest AIC is in bold and includes abrupt monthly changes in CPUE. The models including month with four or five harmonics also have non-zero AIC weights.

Using Bootstrapping for Pre and Post Breakpoint Comparisons

In order to explore potential causes for the breakpoint in CPUE based on fishing effort and catch rates, a selection of key variables have been chosen to explore pre and post breakpoint behaviour. These comparisons of pre and post breakpoint means are carried out using independent two sample t-tests while the corresponding 95% confidence intervals are calculated using bootstrapping.

Variable	Pre-BP mean and 95% CI	Post-BP mean and 95% CI	P-value
CPUE (log scale) (4sf)	6.631 (6.626, 6.637)	4.639 (4.636, 4.613)	2.2 e -16
CPUE (kg/h) (0dp)	10842 (10771, 10914)	4010 (3985, 4034)	2.2 e -16
Catch - broken up by species (kg) (0dp)	2883 (2867, 2899)	1009 (1004, 1014)	2.2 e -16
Catch – total catch per event (kg) (0dp)	6633 (6596, 6670)	4043 (4024, 4062)	2.2 e -16
Fishing duration (hr) (3dp)	0.531 (0.530, 0.533)	0.586 (0.585, 0.588)	2.2 e -16
Trawl distance (km) (2dp)	3.20 (3.19, 3.21)	3.33 (3.32, 3.34)	2.2 e -16
Surface temperature (°C) (2dp)	9.78 (9.78, 9.78)	9.83 (9.83, 9.84)	2.2 e -16

Table 19 A comparison of the means of a selection of relevant explanatory variables pre and post the 2004 breakpoint including 95% confidence intervals and the corresponding p-values from the independent two sample t-tests. This comparison was performed using bootstrapping with a bootstrap sample of 1000. The variable units and degree of rounding for each of the means and confidence intervals are given alongside the variable name in the first column.

CPUE and catch (both total catch and catch by species) experienced significant decreases post the breakpoint in 2004. On the other hand, fishing duration and trawl distance which relate to fishing effort experienced significant increases post breakpoint. The environmental variable surface temperature also experienced a significant increase post breakpoint but further research is required to determine whether this has any direct influence over the orange roughy stock or catch rates.

Locations of trawls either side of the breakpoint were investigated by plotting approximate trawl trajectories using the latitude and longitude coordinates for the start and end of each event. The purpose of this was to investigate the presence of any spatial patterns in fishing effort which could inform the breakpoint. While it is unrealistic to assume the vessels travelled in a straight line from start to end point, the directed line segments (Figure 24) give an indication of the spatial distribution of catches. The clusters give an indication of locations of the geographical features such as seamounts and valleys where aggregations of orange roughy are found. The overlap in fishing locations either side of the breakpoint suggests that the same locations were being targeted for the duration of the fishery.

SUMMARY:

Overall decreases in CPUE and catch, and increases in effort post breakpoint suggest that 2004 may be the beginning of a period of overexploitation which led to the Priceless orange roughy stock becoming both biologically and economically unsustainable. Eventually, this resulted in a voluntary withdrawal of effort from the region and a redirection of effort towards more lucrative stocks in the ORH3B region.

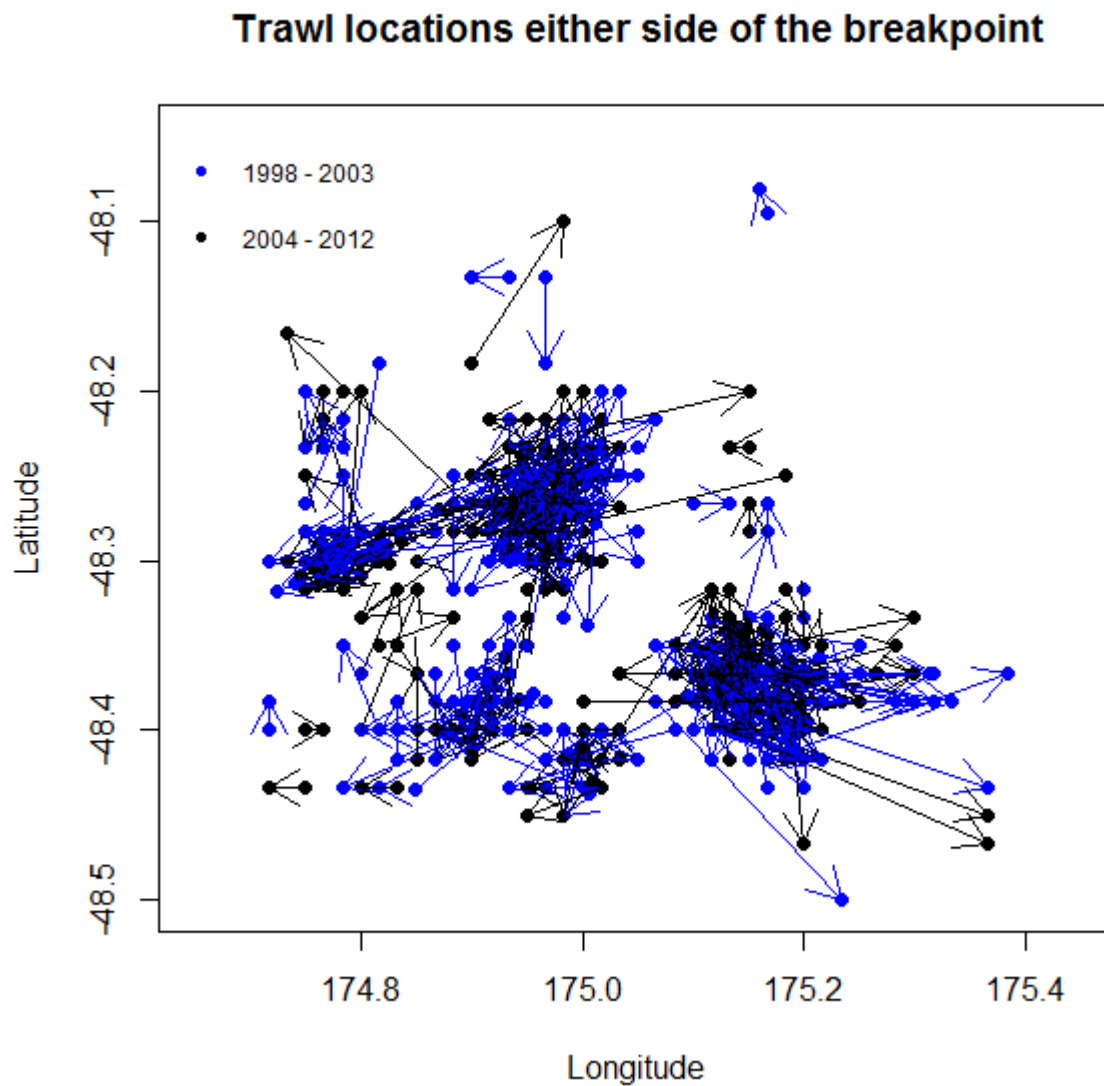


Figure 24 A plot of trawl locations either side of the 2004 CPUE breakpoint. Start and end locations of fishing events have been plotted and joined with directed line segments. The clustering of trawl locations enables identification of the spatial features of the region such as valleys and seamounts. These features correspond to spatial distributions of orange roughy aggregations. The trawl locations have a similar spatial distribution either side of the breakpoint which rules out location as a possible explanation for the breakpoint.

Model Cross Validation and the Pursuit of the Best Model for CPUE

Cross validation was conducted in order to assess the prediction abilities of the LMs for CPUE. Given the large number of models and variables explored, cross validation was performed on a selection of the key models.

Up to this point, the models which best describe CPUE have been identified based on AIC comparisons. The AIC value corresponding to each of the key models is listed alongside the cross validated mean absolute error (MAE) and root mean squared error (RMSE) to allow for a full comparison of the descriptive and predictive abilities associated with each model (Table 20). The errors have been rounded to two decimal places and are arranged in descending order with respect to the MAE.

Cross validation of the CPUE models revealed that the models with the lowest AIC values are not necessarily the models with the best predictive abilities. The six models with the lowest AIC values have the highest prediction errors. Despite having the lowest AICs, these models fail to predict CPUE with the level of accuracy of the remaining models in the table. While the variable surface temperature appears to describe CPUE significantly better than any of the other explanatory variables included here, with respect to the AIC, this does not translate well to predictive ability. The model with the overall lowest AIC has the highest prediction errors (Table 20).

Model	Model AIC	MAE (2dp)	RMSE (2dp)
three_species * month_F * year_F	16102	1.29	1.79
three_species * month_N3 * year_F	16093	1.31	<u>1.95</u>
three_species * month_N3 * year_BP	16582	1.37	1.85
three_species * month_F * year_BP	16553	1.40	<u>1.90</u>
three_species * month_N5 * year_F	16102	1.41	1.89
three_species * month_N * year_BP	17013	1.43	1.93
three_species * month_N4 * year_BP	16556	1.44	1.93
three_species * month_N5 * year_BP	16564	1.44	1.93
three_species * month_N5 * year_N	16914	1.46	1.95
three_species * month_N * year_N	17335	1.48	2.00
surface_temp	13654	2.23	2.67
three_species * month_N * year_N + surface_temp	11893	2.45	3.16
three_species * surface_temp	11900	2.50	3.18
three_species * month_N * year_BP + surface_temp	11693	2.54	3.25
three_species * month_F * year_F + surface_temp	11027	2.65	3.45
three_species * month_F * year_F * surface_temp	10869	3.76	12.19

Table 20 Summary of results from 10-fold cross validation of the key LMs for CPUE. The errors calculated as part of the cross validation are the MAE and the RMSE. The table has been arranged with the MAE in descending order. Two of the RMSEs are larger than those below them in the table which indicates increased variance of the predictions. These two errors are underlined. The inclusion of model AIC values allows for an overall comparison of model performance. Results for models including month as a categorical variable (month_F) should be treated with caution due to issues with rank deficiency. The overall best model for CPUE can be found in bold and includes species, month with three harmonics and an annual linear trend with a breakpoint in 2004.

Discussion

One of the great challenges in evaluating the scientific literature is the difficulty of distinguishing dead bones transferred from one grave to another from sound estimates based on real data (Pauly, 2010).

Heated Explanations of CPUE

The fact that surface temperature is significantly better at describing CPUE than all of the models including species and temporal terms is initially surprising. Given that orange roughy and oreo live deep below the surface (based on the bottom depth variable, orange roughy in the Priceless region are found 760 - 1170m below the surface) where temperature fluctuations are minimal, changes in surface temperature initially appear unlikely to have an influence on orange roughy or oreo behaviour or catchability. Surface temperature is also unlikely to affect catchability by way of equipment efficiency.

A possible explanation for the performance of surface temperature is its ability to form a pseudo time series, incorporating a level of seasonality unavailable from the year and month variables. Due to natural climatic fluctuations, the effect of a new season may be felt at a different time to the calendar start date for the season.

While surface temperature seems to be able to take a lot of temporal information into account, it does not explain it all. The addition of year and month to the model compliments the seasonality explained by sea surface temperature.

On top of the potential to represent seasonality, surface temperature may be related to orange roughy habitat distribution. Temperature sensitivity of orange roughy has been linked to the distributions of juveniles and adults, with juveniles demonstrating a preference for warmer waters. The suspected temperature sensitivity of orange roughy has led some fishers to use net sensors to detect small changes in temperature to enable them to target orange roughy in the warmer areas where it is believed they are more abundant (Dunn, Rickard, Sutton, & Doonan, 2009). It is also thought that water temperature may be a better predictor of orange roughy location than water depth (Branch, 2001).

While the model including all of the interactions between species, month with three harmonics, year with abrupt fluctuations and surface temperature is statistically the best model with an AIC of 10862 (Table 13), it may not be the best representation of reality. Including all the interactions, up to four way, between the four variables in the model leads to uncertainty around whether the interpretation of the model is realistic with respect to biology, oceanography and fishing effort.

And the Winner is...

The ability to predict future CPUE is important due to the assumed relationship between CPUE and abundance. Thus, the predictive ability of the models fitted is as important as their descriptive ability. The results from cross validation provide additional criteria for selection of the most appropriate model to both describe and predict CPUE.

The inclusion of the variable surface temperature leads to predictions which are both highly volatile and lacking accuracy. The model with surface temperature as the sole explanatory variable is oddly the best predictive model of those containing surface temperature despite having the highest AIC. Adding other variables and interactions leads to deterioration in predictive ability. As a result of this, models containing surface temperature are considered unreliable and should not be used to describe or predict CPUE. A likely cause of this unreliability is the missing surface temperature data. The division of the data into 10 sets for cross validation, with a potentially uneven distribution of events with missing surface temperature between groups may explain this.

Cross validation errors on any models which assume abrupt fluctuations in CPUE by month (month_F) should be treated with caution. These models lead to a rank deficient fit where the reliability of the prediction errors is unknown. This rank deficiency is a result of the interaction between year and categorical month. This rank deficiency occurs whether year explains CPUE via abrupt fluctuations, a linear trend or a linear trend with a breakpoint. Based on the distribution generating the data, it makes sense to consider the interaction between year and month so allowing CPUE to fluctuate by month is biologically inappropriate.

Over the fifteen year period being analysed, there are 62 months in which no fishing events took place, corresponding to one third of the total months in the time period. There are ten months in which only one or two fishing events take place. Prediction of these events when month is a categorical variable leads to a rank deficient fit. While the number of cases causing rank deficiency are small relative to the total number of events, it is enough to create uncertainty about the usefulness and validity of these models.

Ignoring the issues with rank deficiency, the 'best' model with respect to both AIC value and cross validation errors assumes abrupt changes in CPUE between years and months and a different CPUE pattern for each of the three species groups. The uncertainty surrounding the reliability of this model due to rank deficient predictions makes it a poor choice.

The two models which rank below this model do not assume abrupt changes in CPUE by month, avoiding the issue of rank deficient predictions (Table 20). The first of these includes abrupt changes in CPUE by year and a seasonal pattern consisting of three peaks and troughs. The rank deficiency issue is resolved by assuming a seasonal pattern in CPUE as opposed to monthly fluctuations, thus the predictions from this model can be treated as reliable.

The RMSE of this model is larger than that of the six models below it, despite its lower MAE. This relatively large difference between the MAE and RMSE indicates the presence of some larger prediction errors which leads to increased variance of the predictions.

On top of this, the presence of abrupt changes in yearly CPUE may compromise the ability of the model to be useful beyond the time frame used in model fitting. Treating CPUE as having yearly fluctuations assumes an element of randomness in the average yearly CPUE. While this makes sense based on the level of noise and uncertainty inherent in CPUE data, it does not provide a direct link between abundance and CPUE.

A better choice would be the model highlighted in bold, which models annual CPUE with a linear trend where the slope of the trend changes at the breakpoint in 2004 (Table 20). This model also incorporates species and seasonality consisting of three peaks and troughs. Despite a larger AIC and MAE, this model aligns more closely to biological assumptions relating to CPUE and has less risk of relatively large prediction errors.

Given that the year variable corresponds to calendar year rather than fishing year, assumptions regarding monthly fishing behaviours are an important factor in determining the most appropriate model. A linear trend in CPUE within each calendar year does not match the observed data. Inspection of Figure 25 suggests the presence of a seasonal pattern with respect to month. The number of events per month were grouped over the 15 years included in the dataset; the month of October being the most common month for fishing events in the Priceless region. The reason for this is that the 1st of October is the start of the fishing year. Once fishers have received their ACE for the new fishing year they are often eager to start filling their quota early to minimise competition from other quota holders. The months in the period from June to August have the lowest number of catches. This relates to the timing of this period in relation to the start of the fishing year as well as the overlap with the orange roughy spawning period. Orange roughy form spawning aggregations from June to early August and the best catch rates occur during this period. As a result, fishing effort during the spawning period is often focused on key orange roughy spawning areas especially the Spawning Box which forms part of the East and South Chatham Rise stock.

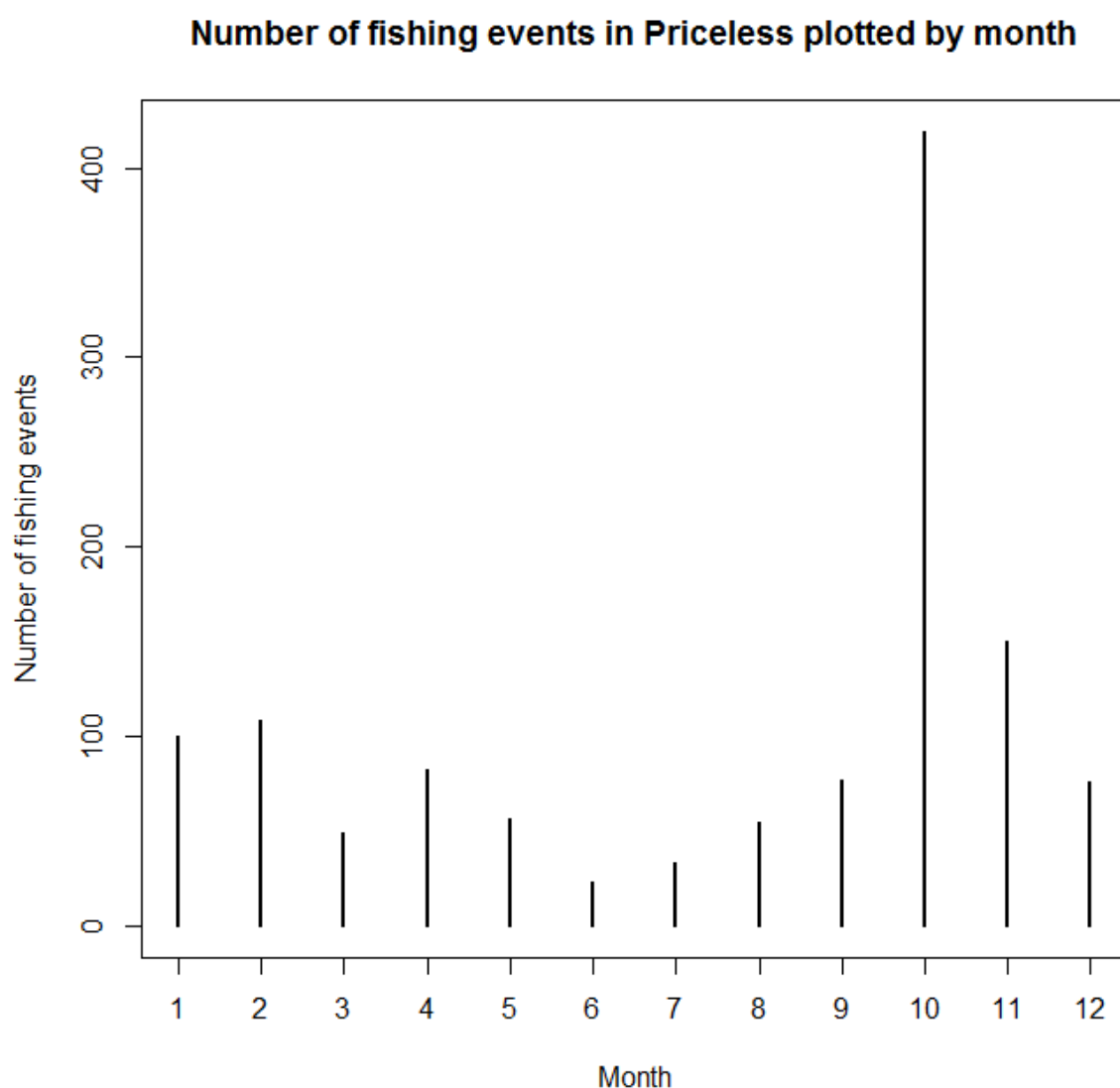


Figure 25 A plot of the number of events in the Priceless region by month over a period of 180 months from January 1998 – December 2012. October is the most popular month for fishing events due to the fact that the fishing year starts on the 1st of October and this corresponds to the start date for the ACEs.

There are three crude peaks and troughs in the number of fishing events with peaks occurring in February, April and October (Figure 25). This data is combined over the full fifteen years of the dataset, but suggests that the assumption of a seasonal pattern in CPUE with three peaks and troughs is sensible based on the observed data.

SUMMARY:

The most useful and adequate model for explaining and predicting CPUE for orange roughy in the Priceless region is, therefore, the model which assumes a linear trend in CPUE with a changing slope in 2004, a seasonal pattern consisting of three peaks and troughs and a different CPUE pattern for each species (Figure 26).

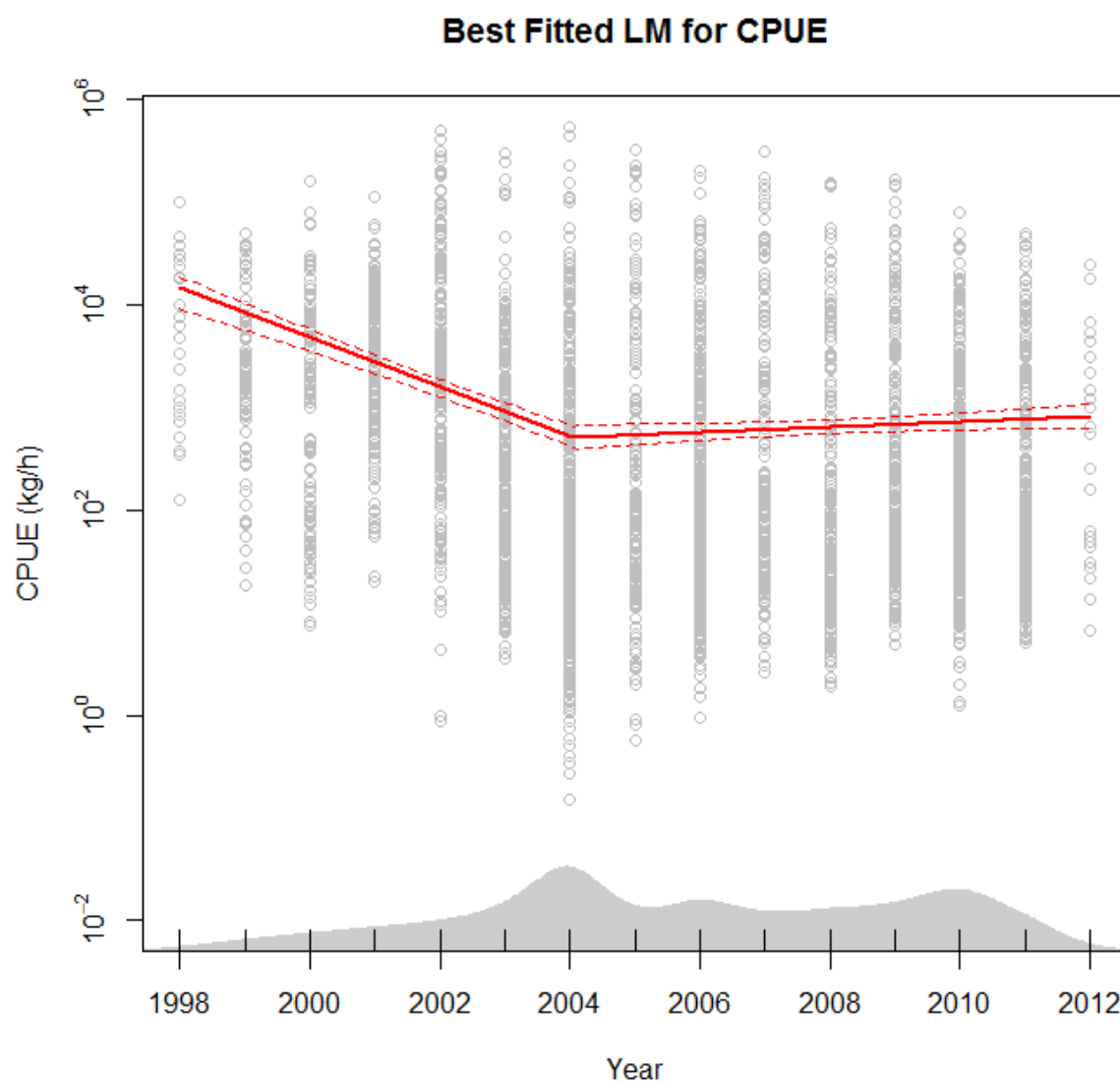


Figure 26 A plot of the fitted vs observed annual CPUE. The fitted model for the best LM of CPUE is plotted (solid red line) with 95% confidence intervals either side (dashed red lines). The fitted model is overlaid on the observed annual CPUE data (empty grey dots) and a contour plot of the number of events per year sits on the x axis to give an indication of relative sample sizes (solid grey contour). The fitted model clearly demonstrates the location of the breakpoint in 2004 and the differing slope either side of this year.

Biological Relevance of a Breakpoint

The presence of a breakpoint in CPUE needs to be interpreted in light of the underlying biology of orange roughy, oreo and the bycatch species involved in the dataset. Visual identification of a breakpoint in the annual CPUE time series hinted at a substantial drop in CPUE around 2003 with the slope being approximately horizontal either side of this drop (Figure 22). From a biological perspective, a substantial drop in abundance is unrealistic. Changes in abundance are expected to approximately follow a negative logistic. The assumption of continuity for orange roughy population dynamics led to the fitting of continuous segmented LMs.

If there was biological evidence in support of a discontinuous breakpoint, it would make sense to fit a selection of discontinuous segmented LMs. Based on the data available and knowledge of orange roughy biology, there is no evidence to suggest an abrupt drop in orange roughy abundance in the Priceless region around 2003-4. For some fisheries, such as the Peruvian anchoveta, abrupt changes in abundance are expected due to their susceptibility to extreme weather events and failed recruitment (Hilborn & Walters, 1992). With a maximum reported age of three years (Whitehead, Nelson, & Wongratana, 1988), the life cycle of the Peruvian anchoveta is at the opposite end of the spectrum to that of the slow growing, centurion orange roughy.

The depth of orange roughy habitat provides a barrier against extreme weather events. Their tendency to spawn sporadically, coupled with their late age at maturity means that the signal of recruitment failures is likely to be indiscernible, especially given the short time series of orange roughy catch relative to their longevity. Depletions in the spawning stock biomass as a result of fishing mortality could take decades to have a discernible effect on recruitment into the subset of the population vulnerable to fishing.

If abrupt changes in orange roughy abundance become apparent, a possible extension of this work would be to consider fitting a discontinuous segmented regression to the CPUE data. Current knowledge of orange roughy population dynamics assert that a continuous breakpoint model is the most appropriate.

While the cause and correct interpretation of the breakpoint remain ambiguous, the general decrease in CPUE is consistent with the inevitable decreases in abundance as a result of fishing pressure.

Practical Usage of a Breakpoint

A significant breakpoint in the annual linear trend for CPUE has been identified. The restrictions placed on the breakpoint due to biological implications have been discussed but there are also practical implications which need to be considered. The breakpoint in annual CPUE was identified after fishing activity in the region had ceased. While identifying a breakpoint retrospectively is an interesting academic exercise, it comes too late to prevent overfishing or population collapse. Had fisheries scientists been aware of this breakpoint in 2004, proactive management may have resulted in a different trajectory for the fishery.

Catch data is highly variable (Plot A in Figure 15 - Figure 18), creating the problem of finding the signal amongst the noise. Cross validation runs that estimated the breakpoint consistently returned the year 2004. Thus, randomly removing 10% of the data, and repeating this process over and over, does not compromise the identification of the breakpoint.

The same may not be true of the removal of one to nine years' worth of data from the end of the time series. A key question to think about is: how long would it take to identify the breakpoint if the model was updated with new data every year? Without an answer to this question, the applicability of using breakpoints to inform management decision or prevent overexploitation is tenuous.

To determine the practical usage of identifying breakpoints in CPUE time series, an obvious extension of this research, would be to create a model which mimics the annual availability of new data.

The breakpoint model involves a segmented linear trend which allows for extrapolation of CPUE outside of the range of the dataset which is an advantage of modelling with a linear trend as opposed to abrupt fluctuations. However, given the potential for breakpoints in the time series, the ability to predict based on a linear trend could result in relatively large prediction errors if a second breakpoint is present.

Does CPUE Reflect Orange Roughy Abundance?

“I think they’re just pissing in the wind.”

– Commentary from Hilborn on Pauly and Zeller’s catch reconstruction research (Cressey, 2015).

Does CPUE reflect the abundance of orange roughy in the Priceless region? The jury is still out on whether catch data and CPUE can or should be used in fisheries science. However, some of the staunchest critics of using catch data admit that it is an essential component of fisheries assessment:

Catch data are a crucial part of any fisheries assessment — it is impossible to calculate the maximum weight of fish that could be harvested sustainably without knowing what is being caught each year. But on their own, catch data cannot answer the question at the heart of fisheries science: how many fish are in the sea?
– Hilborn and Branch (Pauly, Hilborn, & Branch, 2013).

Branch asserts that catch data alone is not sufficient to assess the status of a fishery and that an indication of the actual abundance of a given species is required (Cressey, 2015). While catch data is not sufficient to assess the status of a fishery, it may be necessary to rely on catch trends until more reliable information is made available (Branch, 2001).

Throughout this analysis, CPUE has provided some insights into the plight of orange roughy. The general decrease in CPUE throughout the time series is a placeholder for the inevitable decrease in abundance, however, this relationship is unlikely to be proportional so the true trend in abundance remains unknown.

The analysis has raised more questions than it has answered including the cause(s) of the breakpoint. It is clear, however, that it is possible to identify overfishing through CPUE analysis. Assuming that the CPUE can be calculated for all stocks where the commercial catch data has been collected in the same format between stocks, it could prove to be a useful tool for early detection of overfishing in the orange roughy stocks still being fished. For fish stocks where observer data and fisheries independent surveys are sparse or non-existent, CPUE analysis could prove to be indispensable as a tool to ensure long-term sustainability.

The decrease in CPUE prior to 2004 is steeper than the decrease post-2004. This suggests a possible hyperdepletion relationship between CPUE and abundance. Hyperdepletion between CPUE and abundance for orange roughy has been documented in the literature. Using a Bayesian hierarchical state-space model to analyse the relationship between CPUE and abundance for four orange roughy stocks in New Zealand and Australia yielded an 83% probability of hyperdepletion in favour of hyperstability (Hicks, 2013).

The tendency of orange roughy to form aggregations has the potential to introduce bias in the relationship between CPUE and abundance (Boyer, Kirchner, McAllister, Staby, & Staalesen, 2001). Hyperstability is often seen in CPUE indices of aggregating species.

Without quantitative data on orange roughy abundance, it is difficult to know how well the time series of CPUE reflects abundance. Often the suggestion is made that a time series of abundance estimates should be compared and contrasted with the time series of CPUE. However, this yields an unfortunate catch-22 whereby, if there exists an accurate time series of abundance, analysis of CPUE is superfluous.

The logical next step is to compare the abundance and CPUE time series for orange roughy stocks, where both of these time series exist. Comparisons could then be made between the time series for CPUE between regions which may inform the unknown abundance. This is not without issues either, as there is variation in the biological and environmental conditions between regions which has the potential to invalidate comparisons.

Despite the uncertainty surrounding the use of CPUE to model abundance, a tentative link is favourable to the alternative of being wholly uninformed. And in the words of Pauly:

*The idea that catch data are not useful for
determining the health of fish stocks... is wrong. Dangerously so.
(Pauly, Hilborn, & Branch, 2013).*

A Regional Weigh-in

The Priceless and North Pukaki regions both became victim to overfishing resulting in an informal closure while the Spawning Box and a selection of other orange roughy stocks are still fished to the present day.

The different annual patterns of catch and effort (Figure 15 to Figure 18) indicate that individual stocks can exhibit vastly different responses to fishing pressure. The reasons for these differences may be biological or related to management and fishing practices. These figures indicate that the results from a thorough analysis of one fish stock may not be representative of the situation in other similar stocks.

Given that North Pukaki, Priceless and many other Sub-Antarctic orange roughy stocks are managed as part of the ORH3B stock, which also includes the successful Chatham Rise stocks, it is likely that these stocks were overfished without detection due to management decisions being made at a different spatial scale.

Given that the Spawning Box forms part of the East and South Chatham Rise region, which is one of the three regions currently undergoing Marine Stewardship Council certification (Marine Stewardship Council, 2014), an increasing trend in annual CPUE is expected.

Under the assumption that CPUE has a proportional relationship to orange roughy abundance, the plots of annual CPUE suggest that some orange roughy fisheries are sustainable and some are unsustainable. The increasing trend for the NZ EEZ series suggests that overall New Zealand orange roughy fisheries are being fished at a sustainable level. This observation is a gross oversimplification as the issue of sustainability and orange roughy abundance is much more complicated than a selection of annual CPUE plots.

Future Work

Incorporating Biological Parameters

The pre and post breakpoint analysis focused exclusively on variables relating to catch and effort. This revealed decreases in catch and increases in effort post breakpoint. There are a selection of other variables which are likely to vary either side of the breakpoint which include biological parameters. It would be useful to compare the orange roughy age and length distributions either side of the breakpoint.

Population dynamics, including the distribution of age and length of fish from exploited populations, are subject to change as the fishery develops. The average age and length often decrease over time as fishing mortality removes the older, larger fish from the population (Clark & Tracey, 1994). It is likely that the age and length structures of orange roughy in the Priceless region decreased around the time of the breakpoint. It would be interesting to analyse observer data from the region to investigate this.

Defining the Effort in CPUE

There are many ways of defining CPUE (Dunn, Harley, Doonan, & Bull, 2000). Two definitions of CPUE have been used in this report; kg per event and kg per hour. A possible extension of this work would be to experiment with different definitions of CPUE including kg per kilometre towed or kg per trip which combines all the events occurring in the same trip. It is also interesting to incorporate search time into measures of CPUE. The units for this measure are again kg per hour, but here the time includes the time spent finding a suitable fishing location as well as the time spent with fishing gear in the water. Search time is not included in the orange roughy dataset, but future collection of this variable could prove beneficial for modelling.

Although CPUE measures are usually highly correlated, it is possible that one measure is a better proxy to abundance than the others. Experimentation with a selection of CPUE measures for an orange roughy stock with a known time series of abundance could be useful in determining this relationship.

The CPUE measure can also be modified by changing the units of catch. In this dataset, catch is measured by weight. Changes in the size distribution over time are likely to result in changes to the fishing yield, measured as a count of the number of fish caught. It is possible that analysis of catch at the level of individual fish could provide different insights to catch by weight. Having catch information by weight and as counts would allow for analysis of the average fish weight over time. Count data is not currently included in the fishing database.

The Sub-Antarctic and Beyond

There is also scope for analysing CPUE over different spatial and temporal scales. The exploratory data analysis of the Priceless, North Pukaki, Spawning Box and NZ EEZ regions showed different patterns in the catch and effort variables. It would be interesting to determine whether the best LMs for CPUE in these regions would incorporate the same set of explanatory variables as for CPUE in Priceless.

Analysis of not only North Pukaki but the full Pukaki ranges and other Sub-Antarctic stocks in close proximity to Priceless have the potential to shed light on patterns seen in the Priceless region as well as creating a big picture analysis. All of these stocks, including the Spawning Box and a selection of others are part of the ORH3B QMA. Given that these stocks are managed together, it would be beneficial to analyse the different catch trends occurring within this area, especially given the possibility of migration between these areas.

A comparison of stocks which have collapsed or been subject to overfishing against stocks which are currently being fished at sustainable levels has the potential to reveal patterns which could be used to reduce the risk of overfishing in the future.

Orange roughy and oreo

Oreo catches are included in the dataset used for this analysis. Oreo share many common traits with orange roughy and it is thought that modelling the two species together may lead to new insights. Some basic comparisons and analyses were conducted to explore the relationship between orange rough and oreo, especially with respect to CPUE and abundance. The significance of the species variable in the LMs for CPUE indicate that the two species have distinct CPUE patterns, but there is scope for further investigation of the two species.

Orange Roughy at the End of the Line

Essentially, all models are wrong, but some are useful (Box & Draper, 1987).

The best model for orange roughy CPUE in the Priceless region is wrong (Table 20), but it may be useful. Modelling a complex biological process such as CPUE will always result in some unexplained variation. Despite this, the best model for CPUE can be used to make inferences about orange roughy abundance.

One of the big challenges with modelling orange roughy catch data is that the tendency for a boom and bust fishery often leads to a short time series which limits the ability to create models which will detect potential issues in advance.

The debate continues regarding whether orange roughy fisheries can ever be managed sustainably and whether modelling CPUE can provide adequate estimates of abundance to inform fisheries management in the absence of accurate abundance estimates. In the middle of this debate, one thing remains certain, orange roughy have had a rough time over the past decades and it is our responsibility to ensure that Orange Roughy Grandpa is around for future generations.

Bibliography

- Akaike, H. (1974). A New Look at the Statistical Model Identification. *IEEE Transactions of Automatic Control*, 716-723.
- Anderson, O., & Dunn, M. (2012). *Descriptive analysis of catch and effort data from New Zealand orange roughy fisheries in ORH 1, 2A, 2B, 3A, 3B, 7A, and 7B to the end of the 2008–09 fishing year*. Wellington: Ministry for Primary Industries.
- Andrews, A. H., Tracey, D. M., & Dunn, M. R. (2009). Lead-radium dating of orange roughy (*Hoplostethus atlanticus*): validation of a centurion life span. *Canadian Journal of Fisheries and Aquatic Sciences*, 66: 1130-1140.
- Aranda, M. (2009). Developments in fisheries management in Peru: The new individual vessel quota system for the anchoveta fishery. *Fisheries Research*, 308-312.
- Arlot, S. (2010). A survey of cross-validation procedures for model selection. *Statistics Surveys*, 4: 40-79.
- Arnason, R. (2009). Fisheries management and operations research. *European Journal of Operational Research*, 741-751.
- Australian Fisheries Management Authority. (2014, December 17). *Have your say – Orange Roughy Rebuilding Strategy*. Retrieved from Australian Fisheries Management Authority: <http://www.afma.gov.au/have-your-say-orange-roughy-rebuilding-strategy/>
- Bonfil, R. (2005). The purpose of stock assessment and the objectives of fisheries management. In J. A. Musick, & R. Bonfil, *Management techniques for elasmobranch fisheries* (pp. 6-14). Rome: Food and Agriculture Organization of the United Nations. Retrieved from FAO Corporate Document Repository: <http://www.fao.org/docrep/009/a0212e/A0212E06.htm>
- Box, G., & Draper, N. (1987). *Empirical Model Building and Response Surfaces*. New York: John Wiley & Sons.
- Boyer, D., Kirchner, C., McAllister, M., Staby, A., & Staalesen, I. (2001). The orange roughy fishery of Namibia: lessons to be learned about managing a developing fishery. *South African Journal of Marine Science*, 23: 205-221.
- Braiden, G. (2015, October 31). *Sustainable seafood sales soar as ethical choices of chefs go mainstream*. Retrieved from Herald Scotland: http://www.heraldsotland.com/news/13927624.Sustainable_seafood_sales_soar_as_ethical_choices_of_TV_chefs_go_mainstream/
- Branch, T. (2001). A review of orange roughy *Hoplostethus atlanticus* fisheries, estimation methods, biology and stock structure. *South African Journal of Marine Science*, 23:181-203.
- Brendan, T. O., & Bence, J. R. (2008). Comment: Use of Piecewise Regression Models to Estimate Changing Relationships in Fisheries. *North American Journal of Fisheries Management*, 28:844-846.

- Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? - Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, 7: 1247-1250.
- Clark, M. R., & Tracey, D. M. (1994). Changes in a population of orange roughy, *Hoplostethus atlanticus*, with commercial exploitation on the Challenger Plateau, New Zealand. *Fishery Bulletin*, 92:236-253.
- Clark, P. (2013, October 29). *Delia Smith and Nigella Lawson handed wooden spoon for fishing*. Retrieved from Financial Times: Business & Economy: <http://www.ft.com/intl/cms/s/0/4c99cf1c-40a4-11e3-8775-00144feabdc0.html#axzz40ISmD3yh>
- Clement Group . (2002). *Orange roughy - Hoplostethus atlanticus* . Retrieved from Clement Group : clementgroup.co.nz/clement/species/orh.html
- Climate Prediction Center Internet Team. (2015, November 4). *Cold and Warm Episodes by Season*. Retrieved from NOAA National Weather Service Climate Prediction Center: www.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ensoyears.shtml
- Commonwealth of Australia, Bureau of Meterology. (2016). *Climate Glossary: Southern Oscillation Index*. Retrieved from Australian Government, Bureau of Meterology: www.bom.gov.au/climate/glossay soi/shtml
- Costello, C., Gaines, S. D., & Lynham, J. (2008). Can Catch Shares Prevent Fisheries Collapse? *Science*, 1678-1681.
- Cressey, D. (2015). Eyes on the ocean. *Nature*, 519: 280-282.
- CSGNetwork. (2016). *GPS Latitude and Longitude Converter Version 5.1.4*. Retrieved from CSGNetwork: <http://www.csgnetwork.com/gpscoordconv.html>
- Daniel, H., & Minot, F. (1955). *The Inexhaustible Sea*. London: MacDonald.
- Davison, A., & Hinkley, D. (1997). *Bootstrap Methods and Their Application*. Cambridge University Press.
- Davison, A., & Kuonen, D. (2002). An Introduction to the Bootstrap with Applications in R. *Statistical Computing & Statistical Graphics Newsletter*, 13: 6-11.
- Department of the Environment. (2016). *Hoplostethus atlanticus* . Retrieved from Species Profile and Threats Database, Department of the Environment, Canberra: http://www.environment.gov.au/cgi-bin/sprat/public/publicspecies.pl?taxon_id=68455
- Doonan, I., Horn, P., & Maolagain, C. (2014). *Orange roughy age estimates: Chatham Rise (ORH3B) spawning plumes in 2012, and mid-east coast North Island (ORH2A) fishery from 1989-91 and 2010*. Wellington: Ministry for Primary Industries.
- Dunn, A., Harley, S. J., Doonan, I., & Bull, B. (2000). *Calculation and interpretation of catch-per-unit-effort (CPUE) indices*. Wellington: Ministry of Fisheries.

- Dunn, M. R., Rickard, G. J., Sutton, P. J., & Doonan, I. J. (2009). Nursery grounds of the orange roughy around New Zealand. *ICES Journal of Marine Science*, 66: 871-885.
- Efron, B. (1979a). Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics*, 7: 1-26.
- Efron, B. (1979b). Computers and the Theory of Statistics: Thinking the Unthinkable. *Society for Industrial and Applied Mathematics*, 21: 460-480.
- Efron, B., & Gong, G. (1983). A Leisurely Look at the Bootstrap, the Jackknife, and Cross-Validation. *The American Statistician*, 37:36-48.
- FAO. (2014). *The State of World Fisheries and Aquaculture*. Rome: Food and Agriculture Organization of the United Nations.
- Fenton, G., Short, S., & Ritz, D. (1991). Age determination of orange roughy, *Hoplostethus atlanticus* (Pisces: Trachichthyidae) using $^{210}\text{Pb}/^{226}\text{Ra}$ disequilibria. *Marine Biology*, 109: 197-202.
- Fish Species of New Zealand. (2016). *Orange Roughy*. Retrieved from Fish Species of New Zealand: www.fishspecies.nz/orange-roughy/
- Flowers, J. (2015, November 29). *Popular Fish Fillets that have Absorbed Mercury and are Very Low in Omega-3 Fatty Acids*. Retrieved from Blasting News: <http://uk.blastingnews.com/world/2015/11/popular-fish-fillets-that-have-absorbed-mercury-and-are-very-low-in-omega-3-fatty-acids-00669801.htm>
- Forest & Bird. (2014). *The Best Fish Guide 2013-14*. Retrieved from Forest & Bird: <http://www.forestandbird.org.nz/best-fish-guide-13-14>
- Garcia, S. (1996). The Precautionary Approach to Fisheries and Implication for Fishery Research, Technology and Management: An Updated Review. In *Precautionary approach to fisheries Part 2: Scientific papers*. Rome: Food and Agriculture Organization of the United Nations.
- Gauldie, R., & Cremer, M. (1998). Loss of ^{222}Rn from Otoliths of Orange Roughy, *Hoplostethus atlanticus*, Invalidates Old Ages. *Fisheries Science*, 64: 543-546.
- Gauldie, R., & Romanek, C. (1998). Orange roughy otolith growth rates: a direct experimental test of the Romanek-Gauldie otolith model. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, 120: 649-653.
- GoodFishBadFish. (2011, January 16). *Orange Roughy*. Retrieved from GoodFishBadFish Seafood and Sustainability: goodfishbadfish.com.au/?s=orange+roughy&post_type=fish
- Hicks, A. C. (2013). *The utility of catch-per-unit-effort when assessing and managing long-lived fish stocks*. University of Washington.
- Hilborn, R. (2007). Reinterpreting the State of Fisheries and their Management. *Ecosystems*, 10: 1362-1369.
- Hilborn, R. (2011). Let Us Eat Fish. *The New York Times*.
- Hilborn, R., & Walters, C. J. (1992). *Quantitative Fisheries Stock Assessment: Choice, Dynamics and Uncertainty*. Chapman & Hall.

- Hinton, M. G., & Maunder, M. N. (2003). Methods for standardizing CPUE and how to select among them. *16th Meeting of the Standing Committee on Tuna and Billfish*. Queensland, Australia: Inter-American Tropical Tuna Commission.
- Huxley, T. (1883). Inaugural Meeting of the Fishery Congress. Address by Professor Huxley, F.R.S. *The International Fisheries Exhibition Literature*. London: William Clowes and Sons, Limited.
- Iwamoto, T., Eschmeyer, W., & Alvarado, J. (2010). *Engraulis ringens*. Retrieved from The IUCN Red List of Threatened Species: www.icunredlist.org/details/183775/0
- Jacquet, J. L., & Pauly, D. (2008). Trade secrets: Renaming and mislabeling of seafood. *Marine Policy*, 309-318.
- Kancaid, E. (2015, April 1). *Amazing chart shows the planet's longest-living animals*. Retrieved from Business Insider Australia: <http://www.businessinsider.com.au/animals-that-live-the-longest-2015-3>
- King, M. (2007). *Fisheries Biology, Assessment and Management Second Edition*. Blackwell Publishing.
- Kohavi, R. (1995). A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. *International Joint Conference on Artificial Intelligence*.
- Lack, M., Short, K., & Willock, A. (2003). *Managing risk and uncertainty in deep-sea fisheries: lessons from Orange Roughy*. TRAFFIC Oceania and WWF Endangered Seas Programme.
- Lock, K., & Leslie, S. (2007). *New Zealand's Quota Management System: A History of the First 20 Years*. Wellington: Motu Economic and Public Policy Research Trust.
- Love, P. (2010). *While Stocks Last?* OECD.
- Maceina, M. J. (2007). Use of Piecewise Nonlinear Models to Estimate Variable Size-Related Mortality Rates. *North American Journal of Fisheries Management*, 27:971-977.
- Mackay, K. A. (2005, 12 8). *Database documentation: fish_ce*. Retrieved from NIWA Fisheries Data Management Database Documentation Series: http://www.fish.govt.nz/NR/ronlyres/962CFEA5-1A44-4FD0-A45F-CBC5D17DAF39/1471/researchdatabasedocumentationfish_ce.pdf
- Malakoff, D. (2010, September 2). *War & Fish*. Retrieved from University of Washington Conservation: <http://conservationmagazine.org/2010/09/war-fish/>
- Mansfield, B. (2012, November 9). *Law of the Sea*. Retrieved from Te Ara - the Encyclopedia of New Zealand: <http://www.TeAra.govt.nz/mi/law-of-the-sea/sources>
- Marcus, E. (2011). *Orange Roughy by Ellen Marcus*. Retrieved from Redbubble: www.redbubble.com/people/ellemrcs/works/7800723-orange-roughy
- Marine Conservation Society. (2016). *Good Fish Guide Fish finder: Orange Roughy*. Retrieved from Marine Conservation Society: <http://www.goodfishguide.org/fishfinder?fish=hoplostethus+atlanticus>

- Marine Stewardship Council. (2014, May 30). *MSC welcomes New Zealand orange roughy's brave move to enter assessment*. Retrieved from Marine Stewardship Council:
<https://www.msc.org/newsroom/news/msc-welcomes-new-zealand-orange-roughys-brave-move-to-enter-assessment>
- Maunder, M. N., & Punt, A. E. (2004). Standardizing catch and effort data: a review of recent approaches. *Fisheries Research*, 70: 141-159.
- Melville, H. (1851). *Moby-Dick*. New York: Harper & Brothers.
- Ministry for Primary Industries. (2010). *Orange roughy Fisheries Plan*. Ministry for Primary Industries.
- Ministry for Primary Industries. (2013). *The Status of New Zealand's Fisheries 2013*. Ministry for Primary Industries.
- Ministry for Primary Industries. (2014). *Fisheries Assessment Plenary, May 2014*. Wellington: Ministry for Primary Industries.
- Ministry of Fisheries. (2008). *Harvest Strategy Standard for New Zealand Fisheries*. Wellington: Ministry of Fisheries.
- Ministry of Fisheries. (2008, July). *Orange Roughy (ORH 3B)*. Retrieved from Ministry of Fisheries:
http://www.fish.govt.nz/NR/rdonlyres/34ADE115-9D65-48E1-93E1-7FB15CCEE478/0/ipp_07_08_orh3b.pdf
- Morgan, K. L. (2012, 9 18). *Confidence Intervals: Bootstrap Distribution*. Retrieved from Department of Statistical Science - Duke University:
<https://www2.stat.duke.edu/courses/Fall12/sta101.002/Sec3-34.pdf>
- Mosteller, F., & Tukey, J. W. (1977). *Data Analysis and regression: a second course in statistics*. Addison-Wesley Series in Behavioural Science: Quantitative Methods.
- MRAG Americas. (2015, December 14). *New Zealand Orange Roughy Fishery Stakeholder Notification: Revised Timeline*. Retrieved from Marine Stewardship Council:
https://www.msc.org/track-a-fishery/fisheries-in-the-program/in-assessment/pacific/new_zealand_orange_roughy/assessment-downloads-1/20151217_REV_TLINE_ROU462.pdf
- Muggeo, V. M. (2008). segmented: An R Package to Fit Regression Models with Broken-Line Relationships. *R News*, 8: 20-25.
- New Zealand Government. (2016). *Fisheries Act 1996*. Wellington: New Zealand Government.
- NOAA. (2016, February 8). *Southern Oscillation Index (SOI)*. Retrieved from National Oceanic and Atmospheric Administration:
<https://www.ncdc.noaa.gov/teleconnections/enso/indicators/soi/>
- OECD. (2015). *Country Note on National Fisheries Management -- New Zealand*. Retrieved from OECD: www.oecd.org/newzealand/34430857.pdf

- Orloff, J., & Bloom, J. (2014, May 18). *Bootstrap confidence intervals*. Retrieved from MIT Introduction to Probability and Statistics: http://ocw.mit.edu/courses/mathematics/18-05-introduction-to-probability-and-statistics-spring-2014/readings/MIT18_05S14_Reading24.pdf
- Pauly, D. (2009). *Aquacalypse Now*. *New Republic*.
- Pauly, D. (2010). *5 Easy Pieces: The Impact of Fisheries on Marine Ecosystems*. Island Press.
- Pauly, D., & Zeller, D. (2016). Catch reconstructions reveal that global marine fisheries catches are higher than reported and declining. *Nature Communications*, 7:10244.
- Pauly, D., Hilborn, R., & Branch, T. A. (2013). Fisheries: Does catch reflect abundance? *Nature*, 303-306.
- Picard, R. R., & Cook, R. D. (1984). Cross-Validation of Regression Models. *Journal of the American Statistical Association*, 70: 575-583.
- Pollock, D. (2011). *Topics in Time-Series Analysis: Statistical Fourier Analysis*. Retrieved from University of Leicester: www.le.ac.uk/users/dsgp1/COURSES/LEIMETZ/SPECTRAL.pdf
- Research Data and Reporting Group. (2010, May 18). *Warehou Database Documentation*. Retrieved from Ministry of Primary Industries: <http://www.fish.govt.nz/en-nz/Research+Services/Research+Database+Documentation/Warehou+Database+Documentation.htm>
- Roberts, C. M. (2007). *The Unnatural History of the Sea*. Washington DC: Island Press / Shearwater Books.
- Sanchirico, J. N., & Wilen, J. E. (2007). Global marine fisheries resources: Statis and prospects. *International Journal of Global Environmental Issues*, 106-118.
- Schrope, M. (2010). What's the Catch? *Nature*, 540-542.
- Shao, J. (1993). Linear Model Selection by Cross-Validation. *Journal of the American Statistical Association*, 88: 486-494.
- Shmueli, G. (2010). To Explain or to Predict? *Statistical Science*, 25.3:289-310.
- Sissenwine, M. P., & Mace, P. M. (1992). ITQs in New Zealand: The era of fixed quota in perpetuity. *Fishery Bulletin*, 90:147-160.
- Smart, A. (Director). (2005). *Patrick SmartPants* [Motion Picture].
- Smith, T. (1988). Stock Assessment methods: the first fifty years. In J. Gulland, *Fish Population Dynamics* (pp. 1-28). Avon: John Wiley & Sons, Ltd. .
- Smithsonian National Museum of Natural History. (2015). *Rough Going for Orange Roughy*. Retrieved from Ocean Portal Find Your Blue: ocean.si.edu/ocean-news/rough-going-orange-roughy
- Stevens, C., & O'Callaghan, J. (2015). When the holiday is over: being clever in New Zealand's maine domain. *Journal of the Royal Society of New Zealand*, 45:2, 89-94.

- Stewart, J., & Callagher, P. (2011). Quota concentration in the New Zealand fishery: Annual catch entitlement and the small fisher. *Marine Policy*, 631-646.
- Tesfamichael, D., Pitcher, T. J., & Pauly, D. (2014). Assessing Changes in Fisheries Using Fishers' Knowledge to Generate Long Time Series of Catch Rates: a Case Study from the Red Sea. *Ecology and Society*, 19:18-30.
- Toms, J. D., & Lesperance, M. L. (2003). Piecewise Regression: A Tool for Identifying Ecological Thresholds. *Ecology*, 84: 2034-2041.
- Tracey, D., & Horn, P. (1999). Background and review of ageing orange roughy (*Hoplostethus atlanticus*, Trachichthyidae) from New Zealand and elsewhere. *New Zealand Journal of Marine and Freshwater Research*, 33: 67-86.
- Trenor, C. (2011, March 30). *4 Fish We Should Never Eat*. Retrieved from Alternet: www.alternet.org/story/150407/4_fish_we_should_never_eat
- United Nations, O. (2013). *United Nations Convention on the Law of the Sea - Part V*. Retrieved from United Nations Oceans and Law of the Sea: http://www.un.org/depts/los/convention_agreements/texts/unclos/part5.htm
- Verran, J. A., & Ferketich, S. L. (1987). Testing Linear Model Assumptions: Residual Analysis. *Nursing Research*, 36: 127-129.
- Wagenmakers, E.-J., & Farrell, S. (2004). AIC model selection using Akaike weights. *Psychonomic Bulletin & Review*, 11: 192-196.
- Walrond, C. (2012, November 9). *Fishing Industry*. Retrieved from Te Ara - the Encyclopedia of New Zealand: <http://www.TeAra.govt.nz/en/fishing-industry/print>
- Weeber, B., Thomas, K., & Dorey, C. (2010). *Species Fact Sheet New Zealand orange roughy: Hoplostethus atlanticus*. Greenpeace.
- Whitehead, P., Nelson, G., & Wongratana, T. (1988). FAO Species Catalogue. Vol. 7. Clupeoid fishes of the world (Suborder Clupeoidei). An annotated and illustrated catalogue of the herrings, sardines, pilchards, sprats, shads, anchovies and wolf-herrings. *FAO Fisheries Synopsis*, 125:305-579.
- Worm, B., Barbier, E., Beaumont, N., Duffy, J., Folke, C., Halpern, B., . . . Watson, R. (2006). Impacts of biodiversity loss on ocean ecosystem services. *Science*, 314: 787-90.
- Worm, B., Hilborn, R., Baum, J. K., Branch, T. A., Collie, J. S., Costello, C., . . . Minto. (2009). Rebuilding global fisheries. *Science*, 325: 578-585.
- Wratt, D., Basher, R., Mullan, B., & Renwick, J. (2013). *El Niño and climate forecasting*. Retrieved from NIWA Taihoro Nukurangi: https://www.niwa.co.nz/our-science/climate/information-and-resources/clivar/el_nino
- Zhang, P. (1993). Model Selection Via Multifold Cross Validation. *The Annals of Statistics*, 21: 299-313.